AFRL-OSR-VA-TR-2013-0629

ESTIMATION AND CONTROL FOR LINEAR SYSTEMS WITH ADDITIVE CAUCHY NOISE

JASON SPEYER

REGENTS OF THE UNIV OF CALIFORNIA

12/17/2013
Final Report

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
12-17-2013	FINAL	9/30/2010-9/29/2013

4. TITLE AND SUBTITLE

ESTIMATION AND CONTROL FOR LINEAR SYSTEMS WITH ADDITIVE CAUCHY NOISE

5a. CONTRACT NUMBER

FA9550-10-1-0570

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Jason L Speyer

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of California, Los Angeles

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

USAF/ Office of Scientific Research
Elaine A Ubas 7035888391
CODE FA9550

10. SPONSOR/MONITOR'S ACRONYM(S)

USAF

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A - Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A new class of scalar and vector-state estimators and stochastic controllers for linear dynamic systems with additive Cauchy process and measurement noises has been developed. The Kalman filter and the linear-quadratic-Gaussian controller have been the main estimation and control
paradigms in modern engineering. However, many practical system noises, such as radar glint, are better described by heavy tailed probability density functions (pdf). Although the Cauchy pdf has an infinite variance, the conditional density of a Cauchy random variable, given a linear measurement with an additive Cauchy noise, has a conditional mean and a finite conditional variance, both being functions of the measurement.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Jason Speyer
					19b. TELEPHONE NUMBER *(include area code)* 310-206-4451

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

AFOSR Grant # FA9550-10-1-0570
Estimation and Control for Linear Systems with Additive Cauchy Noise

Final Report

Air Force Office of Scientific Research

Dr. Fariba Fahroo, Program Manager, Dynamics and Controls

Jason L. Speyer
Mechanical and Aerospace Engineering Department
University of California, Los Angeles
Los Angeles, CA 90095-1597

A new class of scalar and vector-state estimators and stochastic controllers for linear dynamic systems with additive Cauchy process and measurement noises has been developed. The Kalman filter and the linear-quadratic-Gaussian controller have been the main estimation and control paradigms in modern engineering. However, many practical system noises, such as radar glint, are better described by heavy tailed probability density functions (pdf). Although the Cauchy pdf has an infinite variance, the ***conditional*** density of a Cauchy random variable, given a linear measurement with an additive Cauchy noise, has a conditional mean and a finite conditional variance, both being functions of the measurement. Over the last three years, a theory of estimation and stochastic control has been developed for the vector state linear dynamic system. The methodology for ***scalar*** state systems entailed propagation of the conditional pdf, while the ***vector*** state case was addressed by developing a recursion for the analytic propagation of the ***character function*** of the unnormalized conditional pdf (ucpdf). Through a spectral transformation, the character function of the ucpdf is used explicitly in the development of stochastic controllers for vector-state systems.

Contents

1 Introduction

In many engineering, economic, telecommunications, and science applications the underlying random processes or noises have significant volatility, which are not captured by Gaussian distributions [1]. Rather than light-tailed Gaussian distributions, heavy-tailed distributions have been shown to better represent these volatile random fluctuations. Examples from the practical engineering world include radar and sonar sensor noises [2], air turbulent environment noise [3], and adversarial motion. Our objective is to develop estimation and stochastic control techniques for linear dynamic systems with heavy-tail distributed noises, which are in the class of symmetric alpha-stable ($S\alpha S$) distributions [4]. In their simplified form, $S\alpha S$ distributions of scalar random variables are characterized by their characteristic function expressed as $\varphi(v) = e^{-\sigma^\alpha |v|^\alpha}$, where σ and α are positive parameters and v is the spectral variable. In this class, $\alpha = 0.5, 1, 2$ yield the Lévy, Cauchy and Gaussian distributions, respectively. For $\alpha \in (0, 2)$, all the densities have infinite variance.

In detection of a radar signal in clutter, the in-phase component of radar clutter time series agrees extremely well with a $S\alpha S$ probability density function (pdf) with $\alpha = 1.7$ [5]. For $\alpha \in [1, 2]$ a maximum likelihood Cauchy detector, which is in the class of myriad filters [6], exhibited performance that is very close to the Cramer-Rao bound, whereas a maximum likelihood Gaussian detector deviated significantly as α varied from 2 to 1. Although numerically intensive, the myriad filters or detectors, based on a cost criterion derived from the α-stable pdfs, show significant improvement in detecting a signal in heavy-tailed noise over Gaussian detectors [7]. Similar performance was observed when processing data in a multi-user communication network [8]. The shortcomings of the Kalman filter and the linear-quadratic-Gaussian (LQG) controller when processing non-Gaussian noise data in the context of aircraft navigation [9] and radar glint [10] suggest that new filters and controllers for heavy-tailed pdfs are required.

The robustness of Cauchy detectors discussed above, together with the deficiencies of the standard Gaussian estimators and controllers when exposed to impulsive noises, motivated the derivation of a sequential non-linear estimator for scalar linear dynamic systems with additive Cauchy process and measurement noises [11, 12]. Although Cauchy noises do not have a well defined first moment and have infinite second and higher moments [13], the conditional pdf (cpdf) of the system state given the measurement history was determined analytically and shown to have well defined, finite conditional first and second order moments [11].

Over the last three years, we have developed the theoretical basis for constructing minimum variance estimators and stochastic controllers for this class of systems. In our initial results for scalar dynamic systems, the conditional pdf of the scalar state given the measurement sequence

could be generated directly [11]. Partial fractions were used to update the conditional pdf at each measurement time. However, this methodology does not generalize when addressing dynamic systems with a vector-state. In [12] a scalar estimator was derived by generating the characteristic function of the unnormalized cpdf (ucpdf) in a recursive scheme. Although the two scalar estimator formulations are similar in form, the characteristic function approach is somewhat simpler than the scheme in [11]. It allows for a stronger result regarding the decay of the estimator parameters with time. Moreover, it was shown that this approach can be generalized to the multi-variable case [14, Appendix A], [15, Appendix B] and for the special two-state system [16], [17, Appendix C].

As shown in [14, Appendix A], the characteristic function of the ucpdf is composed of a sum of terms. Each term has the form of a coefficient, which is a function of the sign of a linear function of the spectral variables, multiplied by an exponential whose argument is a sum of absolute values of the same linear functions of the spectral vector. This analytic form persists through measurement updates and dynamic propagation. Since the characteristic function of the ucpdf is shown to be twice continuously differentiable [14, Appendix A], the conditional mean and the conditional variance are determined by evaluating the characteristic function and its first two differentials as the magnitude of the spectral vector goes to zero.

Although a dynamic programming solution to the stochastic control problem does not currently appear tractable, a stochastic controller has been obtained for the model predictive or open-loop feedback formulation in the scalar case [18] and [19, Appendix D], and extended to the vector case [20, Appendix E], [21, Appendix F]. To ensure that the unconditional expectation of the cost function exists and remains finite for Cauchy uncertainties, and that the conditional expectation of the cost function can be determined in closed form, a new cost function is proposed. This cost function is composed of products of penalty functions on the state and control that are in the form of the Cauchy distribution. Although the conditional expectation of the cost function can be determined from the conditional pdf in the scalar problem, for the vector-state control problem Parseval's Identity is used to transform the conditional expectation of the cost function to spectral domain. Then, the characteristic function of the ucpdf is used directly to determined the conditional expectation of the cost function in closed form [20, Appendix E], [21, Appendix F]. Results on the performance of this stochastic Cauchy controller strongly indicate that it handles outliers in the measurements dramatically well compared to Gaussian controllers. We present our current results for the vector discrete-time Cauchy filter and stochastic controller in section 2.

2 Current Results on Cauchy Filters and Controllers

Over the last three years we have obtained impressive results for vector-state Cauchy estimation, which are detailed in [11, 12, 14 – 17, Appendices A, B & C], are summarized in section 2.1 and our novel control results, which are detailed in [18 – 21, Appedices D, E & F], are summarized in section 2.2.

2.1 Formulation and Derivation of the Cauchy Estimator

The single-input single-output multivariable linear dynamic system[1] is

$$x_{k+1} = \Phi x_k + \Gamma w_k, \quad z_k = H x_k + v_k, \tag{1}$$

where the state vector $x_k \in \mathrm{R}^n$, scalar measurement z_k, and known matrices $\Phi \in \mathrm{R}^{n\times n}$, $\Gamma \in \mathrm{R}^{n\times 1}$, and $H \in \mathrm{R}^{1\times n}$. The noise inputs are assumed to be independent with know Cauchy pdf. Specifically, w_k is assumed to be Cauchy distributed with a zero median and a scaling parameter $\beta > 0$. Similarly, v_k has a Cauchy pdf with a median of zero and a scaling parameter $\gamma > 0$. The characteristic functions of these scalar noises are assumed to be time independent and given by

$$\varphi_W(\bar{v}) = e^{-\beta|\bar{v}|}, \quad \varphi_V(\bar{v}) = e^{-\gamma|\bar{v}|}, \tag{2}$$

where these characteristic functions have a scalar argument $\bar{v}$. The initial conditions at $k = 1$ are also assumed to be independent and Cauchy distributed. Specifically, each i-th element x_{1i} of the initial state vector x_1 has a Cauchy pdf with a zero median and a scaling parameter $\alpha_i > 0$, $i = 1, \ldots, n$. The characteristic function of the joint pdf of the initial conditions, which is a function of a n-dimensional spectral variable $v \in \mathrm{R}^n$, is given by

$$\varphi_{x_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i|v_i|} = \exp\left(-\sum_{i=1}^{n}\alpha_i|v_i|\right) = \exp\left(-\sum_{i=1}^{n} p_i^1|(a_i^1, v)| + j(b_1^1, v)\right). \tag{3}$$

The last form was introduced for notational convenience to be used in the sequel. We used the definitions

$$p_i^1 = \alpha_i, \quad a_i^1 = e_i, \quad i = 1, \ldots, n, \quad b_1^1 = \{0\}_n, \tag{4}$$

where e_i is a n-dimensional i-th unity vector and $\{0\}_n$ is n-dimensional vector of zeros.

[1] The single-input single-output restriction only simplifies the presentation and appears to be easily relaxed.

2.1.1 Characteristic Function for the Un-normalized Conditional pdf

Our goal is to compute the minimum variance estimate of x_k given the measurement measurement history or $\mathbf{z}_k = \{z_1, z_2, \cdots, z_k\}$ [14, Appendix A]. We begin by determining the characteristic function for the un-normalized conditional pdf at $k = 1$, where the conditional pdf at $k = 1$ is

$$f_{X_1|Z_1}(x_1|z_1) = \frac{f_{X_1Z_1}(x_1, z_1)}{f_{Z_1}(z_1)} = \frac{f_{Z_1|X_1}(z_1|x_1)\, f_{X_1}(x_1)}{f_{Z_1}(z_1)} = \frac{f_V(z_1 - Hx_1)\, f_{X_1}(x_1)}{f_{Z_1}(z_1)}. \tag{5}$$

The unnormalized conditional pdf (ucpdf) at $k = 1$ is simply the joint pdf of the vector state and scalar measurement and we will work explicitly with the form $f_V(z_1 - Hx_1)f_{X_1}(x_1)$. The characteristic function of the ucpdf is obtained as

$$\bar{\varphi}_{X_1|Z_1}(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1}(x_1)f_V(z_1 - Hx_1)e^{jx_1^T v}dx_1 = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\varphi_{X_1}(v-\eta)\hat{\varphi}_V(\eta)d\eta. \tag{6}$$

The first integral is a Fourier transform of a product of two functions. Using the dual convolution property, the second integral is a convolution in the v domain between the associated characteristic functions $\varphi_{X_1}(v)$ given in (3) and $\hat{\varphi}_V(\eta)$, the characteristic function of $f_{Z_1|X_1}(z_1|x_1) = f_V(z_1 - Hx_1)$, determined in [14, Appendix A] as

$$\hat{\varphi}_V(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_V(z_1 - Hx_1)e^{jx_1^T v}dx_1 = K(v)\varphi_V\left(-\frac{e_n^T v}{He_n}\right)^{n-1}\prod_{i=1}^{n} 2\pi\delta(e_i P_n v), \tag{7}$$

where $He_n \neq 0$, $P_n = I - \frac{e_n e_n^T}{He_n}$, $K(v) = \exp\left(\frac{jz_1 e_n^T v}{He_n}\right)$, and $\varphi_V\left(-\frac{e_n^T v}{He_n}\right) = \exp\left(-\gamma\frac{e_n^T v}{He_n}\right)$.

Substitution of the $n-1$ delta functions of (7) into (6) reduces the n integrals to one as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi_{X_1}(v - H^T\sigma)\varphi_V(-\sigma)e^{jz_1\sigma}d\sigma = \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi_{X_1}(v - H^T\sigma)\varphi_V(-\sigma)e^{jz_1\sigma}d\sigma, \tag{8}$$

where $\varphi_V(-\sigma) = e^{-\gamma|\sigma|}$.

The convolution integral in (8) is solved in closed form. Here we assume that each element in H is non-zero, i. e., $He_i \neq 0$, $i = \{1, \ldots, n\}$. (See [14, Appendix A] regarding relaxing this condition.) To compute the integral, we define $\rho_i = \alpha_i|He_i|$, $\mu_i = e^T v/He_i$, $i = \{1, \ldots, n\}$,

$\rho_{n+1} = \gamma$, $\mu_0 = -\infty$ and $\mu_{n+1} = \infty$. Hence, the convolution integral is restated and solved as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\sum_{i=1}^{n+1}\rho_i|\mu_i-\sigma|+jz_1\sigma}\,d\sigma = \frac{1}{2\pi}\sum_{R=1}^{n+1} g_i^{1|1}(y_{gi}^{1|1}(v))e^{y_{ei}^{1|1}(v)} \tag{9}$$

where

$$g_i^{1|1}(y_{gi}^{1|1}(v)) = (jz_1+\rho_R+y_{gi}^{1|1}(v))^{-1} - (jz_1-\rho_R+y_{gi}^{1|1}(v))^{-1}, \tag{10}$$

$$y_{gi}^{1|1}(v) = \sum_{\substack{i=1\\ i\neq R}}^{n+1}\rho_i\,\mathrm{sign}(\mu_i-\mu_R), \quad y_{ei}^{1|1}(v) = -\sum_{\substack{i=1\\ i\neq R}}^{n+1}\rho_i|\mu_i-\mu_R| + jz_1\mu_R. \tag{11}$$

As shown in this first update, the coefficient (10) is always a function of sign(·) given by $y_{gi}^{1|1}(v)$ in (11) where the argument of the sign(·) is always found in the argument of the absolute value given by $y_{ei}^{1|1}(v)$ in (11). This is a general property the recursion for the characteristic function of the ucpdf for any update stage time k.

2.1.2 Conditional Mean and Variance at $k = 1$

The conditional mean and variance are computed by evaluating $\bar{\varphi}_{X_1|Z_1}(v)$ and its first two derivatives at $v = \{0\}_n$ or alternatively as $v \to \{0\}_n$. $\bar{\varphi}_{X_1|Z_1}(v)$ can be shown to be twice continuously differentiable [14, 15, Appendices A & B]. Thus, its first two derivatives can be evaluated along a fixed direction $v = E\hat{v}$ while letting $E \to 0$. The pdf of the measurement variable is obtained as

$$f_{Z_1}(z_1) = \bar{\varphi}_{X_1|Z_1}(E\hat{v})_{E=0} = \frac{1}{\pi}\left(\frac{\sum_{R=1}^{n}\alpha_R|h_R|+\gamma}{z_1^2+\left(\sum_{R=1}^{n}\alpha_R|h_R|+\gamma\right)^2}\right). \tag{12}$$

The minimum conditional-variance estimate is given by

$$\hat{x}_1(z_1) = \frac{1}{jf_{Z_1}(z_1)}\left(\frac{\partial\bar{\varphi}_{X_1|Z_1}(E\hat{v})}{E\hat{v}}\right)^{\mathrm{T}}_{E=0} = z_1\frac{[\alpha_1\,\mathrm{sign}(h_1)\cdots\alpha_n\,\mathrm{sign}(h_n)]^{\mathrm{T}}}{\sum_{R=1}^{n}\alpha_R|h_i|+\gamma}, \tag{13}$$

The conditional variance is ($\tilde{x}_1 = x_1 - \hat{x}_1$)

$$E[\tilde{x}_1\tilde{x}_1^T|z_1] = \left(1 + \frac{z_1^2}{\left(\sum_{i=1}^{n}\alpha_i|h_i| + \gamma\right)^2}\right) \begin{bmatrix} \frac{\alpha_1}{|h_1|}\sum_{i=2}^{n}\alpha_i|h_i| + \gamma & \cdots & -\alpha_1\alpha_n\ \mathrm{sign}(h_1)\mathrm{sign}(h_n) \\ \vdots & \ddots & \vdots \\ -\alpha_1\alpha_n\ \mathrm{sign}(h_1)\ \mathrm{sign}(h_n) & \cdots & \frac{\alpha_n}{|h_n|}\sum_{i=1}^{n-1}\alpha_i|h_i| + \gamma \end{bmatrix}. \quad (14)$$

Note that with one measurement, the $n \times n$ conditional variance is bounded and positive definite. Furthermore, the conditional variance is an explicit function of the measurement.

2.1.3 Propagation to $k = 2$ and the Second Measurement Update

The time propagated characteristic function to $k = 2$ is found in [14, 15, Appendices A & B] as

$$\bar{\varphi}_{x_2|z_1}(v) = \bar{\varphi}_{x_1|z_1}(\Phi^T v)e^{-\beta|\Gamma^T v|}. \quad (15)$$

The convolution integral for the second measurement update is

$$\bar{\varphi}_{x_2|z_2}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{\varphi}_{x_2|z_1}(v - H^T\sigma)e^{-\gamma|\sigma|+jz_2\sigma}d\sigma$$

where $\mathbf{z}_2 = \{z_1, z_2\}$. Since $\bar{\varphi}_{x_1|z_1}(v)$ is twice differentiable, then for the linear transformation Φ, $\bar{\varphi}_{x_1|z_1}(\Phi^T v)$ is also continuous. By assuming $H\Gamma \neq 0$, it can be shown that the first two derivatives of $\bar{\varphi}_{x_2|z_2}(v)$ are continuous. If $H\Gamma = 0$, then $e^{-\beta|\Gamma^T v|}$ comes out of the convolution integral. Consequently, $\bar{\varphi}_{x_2|z_2}(v)$ does not have a continuous derivative because the derivative of $e^{-\beta|\Gamma^T v|}$ is only piecewise continuous and thus there is no estimate of x_2

2.1.4 General form of $\bar{\varphi}_{x_k|z_k}(v)$

In general, the convolution integral for the k^{th} measurement update is

$$\bar{\varphi}_{x_k|z_k}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{\varphi}_{x_k|y_{k-1}}(v - H^T\sigma)e^{-\gamma|\sigma|+jz_k\sigma}d\sigma, \quad (16)$$

where the measurement history is $\mathbf{z}_k = \{z_1, \ldots, z_k\}$. It is assumed that $H\Gamma \neq 0$ at each time step and that by induction $\varphi_{x_{k-1}|Y_{k-1}}(v)$ is twice differentiable. Therefore, $\varphi_{x_k|Z_k}(v)$ is twice continuously differentiable. It is expressed in a closed form [14, 15, Appendices A & B] as

$$\bar{\varphi}_{x_k|Z_k}(v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\, \gamma_{gi}(v) \exp\left(y_{ei}^{k|k}(v)\right), \tag{17}$$

where

$$y_{gi}^{k|k}(v) = \sum_{R=1}^{n_{ei}^{k|k}} q_{iR}^{k|k}\, \mathrm{sign}(\langle a_{iR}^{k|k}, v\rangle) \in \mathbb{R}^k, \quad y_{ei}^{k|k}(v) = -\sum_{R=1}^{n_{ei}^{k|k}} p_{iR}^{k|k}\, |\langle a_{iR}^{k|k}, v\rangle| + j\langle b_i^{k|k}, v\rangle. \tag{18}$$

In the above, $q_{iR}^{k|k} \in \mathbb{R}^k$, $p_{iR}^{k|k}$, $a_{iR}^{k|k} \in \mathbb{R}^n$, and $b_i^{k|k} \in \mathbb{R}^n$ are all parameters computed recursively up to time k. This functional form persists at every time step.

Note that this form (17) is also consistent with the characteristic function of the ucpdf for the scalar system [12]

$$\bar{\varphi}_{x_k|Z_k}(v) = \sum_{i=1}^{k+2} a_i(k|k) + jb_i(k|k)\,\mathrm{sign}(v)\; e^{-\omega_i(k|k)|v| + j\sigma_i(k|k)v}. \tag{19}$$

The parameters $a_i(k|k)$ and $b_i(k|k)$ satisfy a linear discrete dynamic equation, from which it is shown that they converge to zero as $k - i$ becomes large. Consequently, these decaying terms can be removed from the sum in (19). This property should generalize for the $g_i^{k|k}\, y_{gi}^{k|k}(v)$ in (17), in which one of the central goals of the proposed project is to prune the sum in (17) such that $n_t^{k|k}$ remains bounded.

There are some simplifications that occur in the two-state estimation problem [16], [17, Appendix C]. A recursion in the terms in the argument of the exponential can be made explicit. A basis B_k for each term in the sum is recursive as

$$B_k = \begin{bmatrix} B_{k-1}\Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \end{bmatrix}, \; B_1 = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}, \; \mathrm{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{20}$$

This basis is related to $a_{iR}^{k|k} \in \mathbb{R}^n$ in the general solution (18) in that every $a_{iR}^{k|k}$ vector is coaligned with a row in B_k. This bases appears minimal and thereby all terms having the same argument of the exponential can be combined. Generalization of this result for the multivariate case will be considered also in the new project.

2.1.5 The Mutlivariable Estimator

As shown in [14, 15, Appendices A & B], to construct the conditional mean and variance, choose $v = E\hat{v}$ where $E > 0$ and $\hat{v}$ is a fixed direction. Assuming the condition that $(a_{iR}^{k|k}, \hat{v}) \neq 0 \ \forall (i, £)$, then

$$\text{sign}((a_{iR}^{k|k}, E\hat{v})) = \text{sign}((a_{iR}^{k|k}, \hat{v})) = s_R^i, \tag{21}$$

$$y_{gi}^{k|k}(E\hat{v}) = \sum_{R=1}^{n_{ei}^{k|k}} q_{iR}^{k|k} s_R^i = y_{gi}^{k|k}(\hat{v}), \ y_{ei}^{k|k}(E\hat{v}) = E(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v}), \tag{22}$$

where s_R^i is a piecewise constant and $(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})$ is a constant. The pdf of the measurement history is determined as

$$f_{Z_k}(\mathbf{z}_k) = \bar{\varphi}_{x_k|Z_k}(E\hat{v})\Big|_{E=0} = \prod_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}^{k|k}(\hat{v})) \ . \tag{23}$$

The conditional mean of the state x_k is

$$\hat{x}_k = E[x_k|\mathbf{z}_k] = \frac{1}{j f_{Z_k}(\mathbf{z}_k)} \left[\frac{\partial(-\bar{x}_{k|Z_k}(E\hat{v}))}{\partial(E\hat{v})}\right]^T\Big|_{E=0} = \frac{1}{j f_{Z_k}(\mathbf{z}_k)} \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}^{k|k}(\cdot)) \, y_{ei}^{k|k}(\cdot), \tag{24}$$

and the second moment

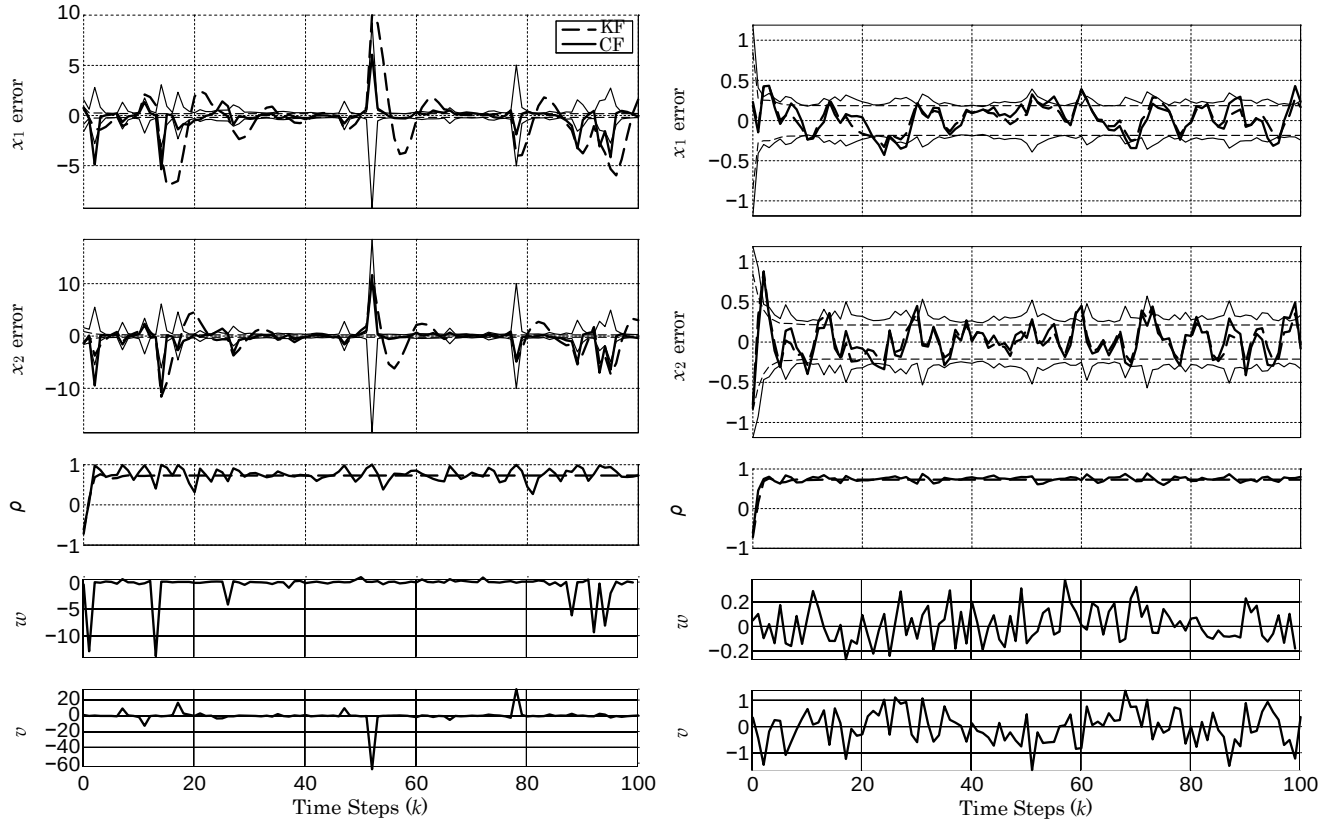
$$E[x_k x_k^T|\mathbf{z}_k] = \frac{1}{j^2 f_{Z_k}(\mathbf{z}_k)} \frac{\partial^2(-\bar{x}_{k|Z_k}(E\hat{v}))}{\partial(E\hat{v})\partial(E\hat{v})^T}\Big|_{E=0} = \frac{-1}{f_{Z_k}(\mathbf{z}_k)} \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}^{k|k}(\hat{v})) \, \bar{y}_{ei}^{k|k}(\hat{v}) \, \bar{y}_{ei}^{k|k}(\hat{v})^T. \tag{25}$$

For the error $\tilde{x}_k = x_k - \hat{x}_k$, the conditional error variance can be evaluated as $E[\tilde{x}_k \tilde{x}_k^T|\mathbf{z}_k] = E[x_k x_k^T|\mathbf{z}_k] - \hat{x}_k \hat{x}_k^T$.

2.1.6 Numerical Examples of the Vector-State Cauchy Estimator

Estimation results for a two-state system with Cauchy and Gaussian noises, related by a least square fit, are shown in Fig. 1. For the two-state system, data can be streamed continuously by using a finite data window, as given in [17, Appendix C]. The error is shown to be in the sixth place, thereby insignificant. In the subfigure 1(a) we see that the Gaussian estimator error deviates from the minimal Cauchy conditional mean error in the Cauchy simulation. Also we note that the actual conditional standard deviation, generated by the Cauchy estimator, fluctuates dramatically with the Cauchy noises. For the Gaussian simulation of Fig. 1(b) both the Cauchy and the Gaussian filters perform almost identically, although the conditional standard deviation *generated*

by the Cauchy estimator slightly upper bounds the actual conditional standard deviation, ***a priori*** computed for the Gaussian estimator. Numerical results for the three-state Cauchy estimator can be found in [14, 15, Appendices A & B].



(a) *Cauchy Noise Simulation* (b) *Gaussian Noise Simulation*

Figure 1: Cauchy and Kalman Estimators for a stable system. Simulation parameters are: $\alpha_1 = \alpha_2 = 0.8$, $\gamma = 0.5$, $\beta = 0.1$, H = [1 1], $\Gamma = [0.5,\ 1]^T$, and the eigenvalues of Φ are $0.8 \pm .55j$. Heavy lines are the state estimate. light lines depict the predicted standard deviation.

2.2 The Stochastic Model Predictor Cauchy Controller

A stochastic optimal control problem is formulated where the conditional performance index, which is the conditional expectation of the cost function, is taken with respect to the Cauchy conditional pdf. However, since only the characteristic function of the conditional pdf is available, a transformation to the spectral vector is required which then leads to a closed-form expression for the conditional expectation of the cost function. In order for the conditional performance index to be analytic, a new cost function is chosen.

The dynamic stochastic system

$$x_{k+1} = \Phi x_k + \Lambda u_k + \Gamma w_k, \quad z_k = \mathrm{H} x_k + v_k \tag{26}$$

with measurement history

$$\mathbf{z}_k = \{z_1, \ldots, z_k\} \tag{27}$$

is the same as (1) except for the addition of the scalar control u_k and known matrix $\Lambda \in \mathrm{R}^n$. To determine a form of the conditional performance index, where the control enters in a more convenient manner, the state vector is decomposed into a dynamic system forced by the control and a dynamic system forced by the underlying random variables. Given this decomposition, we show that the control has to be adaptive to only the σ-algebra generated by the measurement history associated with the stochastic part of the state decomposition.

Consider the linear, discrete-time, scalar stochastic system of (26) with the measurement history given by (27). Let u_k be adaptive to the filtration σ-algebra σ_k generated by the measurement history $\mathbf{z}_k$. Filtration implies that the collection of σ-algebras σ_k have the property that if $j \leq k$, then $\sigma_j \subseteq \sigma_k$ [22]. Therefore, filtration is the evolution of the σ-algebra generated by measurement history through time. Adaptation means that the control is a measurable function of events on this σ-algebra, i.e., this ensures that the control sequence is causal. Now consider the decomposition $x_k = \tilde{x}_k + \bar{x}_k$ where

$$\tilde{x}_k = \Phi \tilde{x}_{k-1} + w_{k-1}, \quad \tilde{z}_k = H\tilde{x}_k + v_k, \tag{28a}$$

$$\bar{x}_k = \Phi \bar{x}_{k-1} + u_{k-1}, \quad \bar{z}_k = H\bar{x}_k. \tag{28b}$$

Here, $\tilde{x}_k$ and $\tilde{z}_k$ are the state and the measurement of the subsystem containing all the underlying random variables, i. e., w_k, v_k, and the initial condition $\tilde{x}_0$, which is Cauchy distributed with zero median. Similarly, $\bar{x}_k$ and $\bar{z}_k$ are the state and measurement of a dynamic system driven by u_k with initial condition $\bar{x}_0$.

The measurement history can be decomposed as $\mathbf{z}_k = \tilde{\mathbf{z}}_k + \bar{\mathbf{z}}_k$ where

$$\tilde{\mathbf{z}}_k = \{\tilde{z}_0, \cdots, \tilde{z}_k\}, \quad \bar{\mathbf{z}}_k = \{\bar{z}_0, \cdots, \bar{z}_k\}. \tag{29}$$

In the following it is shown that the control is measurable on events generated by $\tilde{\mathbf{z}}_k$ only.

Theorem 1. *Consider the filtration $\sigma-$algebra $\tilde{\sigma}_k$ generated by $\tilde{\mathbf{z}}_k$, with the decomposition $\mathbf{z}_k = \tilde{\mathbf{z}}_k + \bar{\mathbf{z}}_k$. For $\tilde{\mathbf{z}}_k \in \tilde{\sigma}_k$ and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, $\bar{\mathbf{z}}_k$ is adapted to $\tilde{\sigma}_{k-1}$ and u_k is adapted to $\tilde{\sigma}_k$.*

Proof. Start with $k = 0$. The initial state is decomposed as $x_0 = \tilde{x}_0 + \bar{x}_0$, where $\bar{x}_0$ is a given

non-random parameter. The measurement decomposes as $z_0 = \tilde{z}_0 + \bar{z}_0$, where $\bar{z}_0 = H\bar{x}_0$ is a given non-random parameter and $\tilde{z}_0 = \tilde{\mathbf{z}}_0 \in \tilde{\sigma}_0$. Then, u_0, which is determined by z_0, is adapted to $\tilde{\sigma}_0$. At $k = 1$, both $\bar{x}_1 = \Phi\bar{x}_0 + u_0$ and $\bar{z}_1 = H\bar{x}_1$ are adapted to $\tilde{\sigma}_0$, and thus $\bar{\mathbf{z}}_1$ is adapted to $\tilde{\sigma}_0$. For the measurement at $k = 1$, $\tilde{z}_1 \in \tilde{\sigma}_1$, $\tilde{\mathbf{z}}_1 \in \tilde{\sigma}_1$, and $\tilde{\sigma}_0 \subset \tilde{\sigma}_1$. Hence, since u_1 is determined by $\mathbf{z}_1 = \tilde{\mathbf{z}}_1 + \bar{\mathbf{z}}_1$, it is adapted to $\tilde{\sigma}_1$. Recursively to any k, $\bar{\mathbf{z}}_k$ is adapted to $\tilde{\sigma}_{k-1}$. With $\tilde{\mathbf{z}}_k \in \tilde{\sigma}_k$, and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, u_k that is determined by $\mathbf{z}_k = \tilde{\mathbf{z}}_k + \bar{\mathbf{z}}_k$ is adapted to $\tilde{\sigma}_k$. □

2.2.1 Cost Criterion for Controller of Cauchy System

Since dynamic programming does not have a tractable solution [18, 19, Appendix D], a multi-step model predictor Cauchy controller is proposed. The maximization for the unconditional expectation of the cost function is

$$J^* = \max_{U_k^{p-1} \in F} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right)\right] = E\left[\max_{U_k^{p-1} \in F} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \mid \tilde{\mathbf{z}}_k\right]\right] = E\left[\max_{U_k^{p-1} \in F} J_{\tilde{z}_k}\right] = E\left[J^*_{\tilde{z}_k}\right], \quad (30)$$

where $\psi\left(X_{k+1}^p, U_k^{p-1}\right)$ is given in (31), the unconditional expectation is assumed to exist and the expectation is nested where the outer expectation is over the measurement history and the inner expectation is over all other random variables. The operations of maximization and expectation have been interchanged [23] and the conditioning on $\tilde{\mathbf{z}}_k$ is justified by Theorem 1. The arguments of the cost function are the projected state and control, given as $X_{k+1}^p := \{x_{k+1}, \ldots, x_p\}$ and $U_k^{p-1} := \{u_k, \ldots, u_{p-1}\}$, $U_k^{p-1} \in F$, where F is the class of piecewise continuous functions adapted to $\tilde{\mathbf{z}}_k$.

The cost function for the Cauchy dynamic system has penalty functions of the Cauchy pdf form as

$$\psi\left(X_{k+1}^p, U_k^{p-1}\right) = \prod_{i=k}^{p-1} f_{u_i}(u_i) \cdot f_{x_{i+1}}(x_{i+1}) = \prod_{i=k}^{p-1} \left[\frac{\zeta_i/\pi}{u_i^2 + \zeta_i^2} \cdot \prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{x_{i+1,r}^2 + \eta_{i+1,r}^2}\right]. \quad (31)$$

To compare the Cauchy controller to a Gaussian controller, an exponential cost function is suggested as

$$\psi\left(X_{k+1}^p, U_k^{p-1}\right) = \prod_{i=k}^{p-1} \left[e^{-\frac{1}{2}r_i u_i^2} \cdot \prod_{r=1}^{n} e^{-\frac{1}{2}q_{i+1,r}x_{i+1,r}^2}\right]. \quad (32)$$

A solution to this LEG problem can be found essentially in [23], but it is explicitly given for the scalar problem in [19, Appendix D].

2.2.2 The Conditional Performance Index in Terms of the Spectral Variables

Since only the characteristic function of the ucpdf is available (See section 2.1.5) and thereby the characteristic function of the conditional pdf, the conditional performance index[2] can be rewritten by applying Parseval's identity directly [20, 21, Appendex E & F] as

$$
\begin{aligned}
J_{\tilde{Z}_k} &= \prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{\sigma_i^2+\zeta_i^2} \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \pounds_{x_p}(\tilde{x}_p+\bar{x}_p)\cdot f_{\tilde{X}_p|\tilde{Z}_k}(\tilde{x}_p|\tilde{z}_k)\, d\tilde{x}_{p,1}\ldots d\tilde{x}_{p,n} \\
&= \frac{1}{(2\pi)^n}\prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{\sigma_i^2+\zeta_i^2} \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} L^*_{x_p}(v)\cdot\varphi_{\tilde{X}_p|\tilde{Z}_k}(v)\, dv_1\ldots dv_n \\
&= \frac{1}{(2\pi)^n}\prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{\sigma_i^2+\zeta_i^2} \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\prod_{r=1}^{n} e^{-\eta_{p,r}|v_r|+j\bar{x}_{p,r}v_r}\,\varphi_{\tilde{X}_p|\tilde{Z}_k}(v)\, dv_1\ldots dv_n, \qquad (33)
\end{aligned}
$$

where L^* refers to the complex conjugate of L. Given the projected form of the characteristic function of the ucpdf (17) and that $\varphi_{\tilde{X}_p|\tilde{Z}_k}(v) = f_{\tilde{Z}}(\tilde{z})\bar{\varphi}_{\tilde{X}_p|\tilde{Z}_k}(v)$, $J_{\tilde{Z}_k}$ can be evaluated in closed form. In the next section, the closed form of the conditional performance index is determined explicitly for the scalar system. For the two-state system, the conditional performance index is found in closed form in [20, 21, Appendex E & F].

2.2.3 Closed-Form Cost Criterion: Scalar State System

The scalar state example considers only a weight on the terminal horizon state x_p, a weight on $u_i\ \forall i \in \{k,\ldots p-1\}$, and the terminal horizon is $m = p-k$. The cost criterion is

$$
J^{*k}_{\tilde{Z}} = \max_{U_k^{p-1}\in F} = \frac{1}{(2\pi)^n}\prod_{i=k}^{p-1}\frac{\zeta_i/\pi}{\sigma_i^2+\zeta_i^2}\int_{-\infty}^{+\infty} e^{-\eta_p|v|+j\bar{x}_p v}\,\varphi_{\tilde{X}_p|\tilde{Z}_k}(v)\, dv, \qquad (34)
$$

where (19) is used to obtain the projected form of the characteristic function of the ucpdf $\bar{\varphi}_{X_p|Z_k}(v)$. The cost criterion, found in closed form as

$$
J_{\tilde{Z}_k} = \prod_{i=k}^{p-1}\frac{1}{\sigma_i^2+\zeta_i^2}\cdot\prod_{p}^{k+2}\sum_{i=1}^{m}\omega_i\frac{a_i(\bar{x}_p+\sigma_i)\eta_p+(b_i-a_i\bar{x}_p)(\eta_p+\omega_i)}{(\bar{x}_p+\sigma_i)^2+(\eta_p+\omega_i)^2}, \qquad (35)
$$

is to be maximized with respect to $U_k^{p-1}\in F$ and subject to the deterministic state as propagated as $\bar{x}_p = \Phi^m\bar{x}_k + \sum_{i=1}^{m}\Phi^{m-1}\Lambda u_{k+i-1}$.

[2] For simplicity we consider only the terminal state in cost function.

2.3 Numerical Examples of the Scalar and Two-State Stochastic Controller

To obtain insight into the properties of the Cauchy stochastic controller, the one-step one-measurement example is first analyzed in section 2.3.1. Next, the multi-step numerical simulation results are given in section 2.3.2, which illustrate the Cauchy controllers behavior in the presence of dominate Cauchy measurement noise and then dominant Cauchy process noise.

2.3.1 One-Step One-Measurement Examples for Scalar and Two-state Systems

The value of the optimal control signal at $k = 0$, i.e., u_0^*, as a function of the first measurement $\tilde{z}_0$, is determined. Specifically, we examine the value of the optimal control signal at $k = 0$, i.e., $\eta_0^*(z_0)$, as a function of the first measurement $\tilde{z}_0$, that varies due to the measurement noise v_0 [19, Appendix D], while considering the one step horizon, i. e., $m = 1$. The parameters for the system and Cauchy signals are first chosen as

$$\Phi = 1,\ H = 1,\ \alpha = 0.1,\ \beta = 0.02,\ \gamma = 0.5,\ \bar{x}_0 = 0. \tag{36}$$

Initially, no penalty is introduced on the control signal, i.e., the term $\zeta_0^2/(u_0^2 + \zeta_0^2)$ is removed from the objective function in (35), i.e. $\zeta_0 \to \infty$ while the state at $k = 1$ is weighted with $\eta_1 = .7$. Therefore, we first examine the case of one step horizon, i.e., $m = 1$. Substituting these parameters into (35), the performance index becomes

$$J_{\tilde{Z}_0}^* = \frac{0.1148(4.1667z_0^2 - 1.0163u_0z_0 + 1)}{(z_0^2 + 0.16)(u_0^2 + 0.6724)} + \frac{0.03416(7.5820z_0^2 + 3.4153u_0z_0 - 1)}{(z_0^2 + 0.16)((u_0 + z_0)^2 + 1.22)} \tag{37}$$

The optimal controller can be obtained by minimizing (37) with respect to u_0. The necessary optimality condition, $\partial J_{\tilde{Z}_0}^* / \partial u_0 = 0$, reduces to finding the roots of the fifth-order polynomial

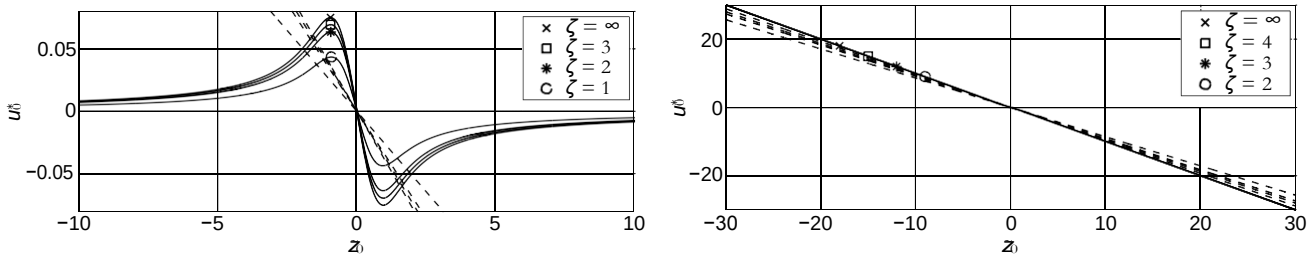
$$l_5u^5 + l_4u^4 + l_3u^3 + l_2u^2 + l_1u + l_0 = 0, \tag{38}$$

where

$$\begin{aligned} &l_5 = 1, \quad l_4 = 3.5z_0, \quad l_3 = (5.2315z_0^2 + 3.6681) \quad l_2 = (3.6806z_0^3 + 6.6305z_0), \\ &l_1 = (0.9491z_0^4 + 3.2124z_0^2 + 2.9623), \quad l_0 = (0.07782z_0^2 + 0.3992)z_0. \end{aligned} \tag{39}$$

This polynomial always has at least one real root. If three roots are real, then there are two local maximum values and the larger of the two gives the optimal control.

Solving the polynomial numerically, the optimal control signal is plotted versus the measurement $\tilde{z}_0$ in Fig. 2(a) for $\zeta_0 = \infty$ as well with weighting $\zeta_0 = 1,\ 2,\ 3$ and $\gamma > \alpha$. Also included is the LEG controller where $u^* = -\hat{x}$, i.e., it is linear in z (see [19, Appendix B] for details of the LEG model predictive controller). The Cauchy controller in Fig. 2(a) for large $\tilde{z}_0$ goes toward zero. This is in sharp contrast with the LEG controller, which remains linear in the measurement. This is a significant difference in behavior between the Cauchy and Gaussian optimal controllers that can be deduced analytically from (38). If $u^*(\tilde{z}_0)$ is finite, the dominant term in (38) as $|\tilde{z}_0| \to \infty$ is $l_1 u^*(\infty)$, or $\lim_{|\tilde{z}_0|\to\infty} u^*(\tilde{z}_0) \to 0$. Therefore, the problem of handling outliers, which occur for the Cauchy pdf, appears to be resolved by the Cauchy controller explicitly, and not in some filter as has been done traditionally. Note that the controller design process explicitly uses the parameters $\gamma > \alpha$, i. e., it should expect more impulsive measurement uncertainty than process uncertainty. If $\gamma < \alpha$, then the Cauchy controller behaves approximately like the LEG linear controller in Fig. 2(b), i. e., it should expect more impulsive process uncertainty than measurement uncertainty. The effect of reducing ζ_0 from ∞ to 5 has no effect on either the Cauchy or LEG controllers, although further reductions do have a small effect as seen in Fig. 2(a) where $\gamma > \alpha$ and approximately linear Cauchy controller for $\gamma < \alpha$ as shown in Fig. 2(b). However, for $\gamma < \alpha$ a homotopy optimization method [19, Appendix B] keeps the optimum value on the ridge that emphasizes the terminal state, rather than emphasizing small control values. This occurs even if eventually the control results from a local optimum and not the global optimum.



(a) $\eta = .7,\ \alpha = 0.1,\ \gamma = 0.5,\ \beta = 0.02,\ \zeta = 2,\ 3,\ 4,\ \infty$ (b) $\eta = .7,\ \alpha = 0.5,\ \gamma = 0.1,\ \beta = 0.02,\ \zeta = 2,\ 3,\ 4,\ \infty$

Figure 2: Scalar Cauchy and Gaussian one-step controller with parameters variations in ζ for $\gamma > \alpha$ left and $\gamma < \alpha$ right.

For a two-step prediction of a two-state stochastic controller in [20, 21, Appendices E & F], all the simulations use the same system dynamics with H = [1 1], Γ^T = [0.5 1], Λ^T = [0.5 1], $m = 2$, and the eigenvalues of Φ are $0.8 \pm 0.55j$. The terminal state weightings are η = [1 1], and when control weightings are used they are $\zeta = 10$. The initial condition's scaling parameters are given by α = [0.8 0.8], and the process and measurement noise parameters β and γ are either 0.5 or 0.1. The first set of examples are shown in Fig. 3. These figures show the applied optimal control input at the first time step given the first measurement. In the two cases presented, all the systems parameters are the same, except in Fig. 3(a) $\gamma > \beta$ (i.e., more measurement than state

uncertainty), and in Fig. 3(b) $\beta > \gamma$ (i.e., more state than measurement uncertainty).

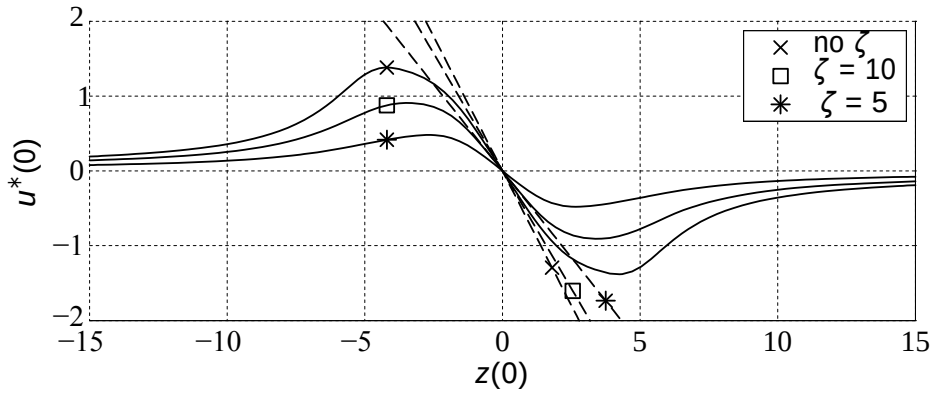
The example in Fig. 3(a) shows that the Cauchy controller is nearly linear for small measurements and reduces its control effort to zero as the measurement deviations become large. This is in contrast to the LEG controller, which is linear and thus responds strongly to large measurement deviations. This behavior in the Cauchy controller occurs when the measurement uncertainty is larger than the state uncertainty. In the opposite case shown in Fig. 3(b), the measurement has less uncertainty than the state. Here, the Cauchy controller's response closely matches that of the LEG in a neighborhood of the origin, and in fact responds even more strongly than the LEG for large measurement deviations.

The three different curves in both of these figures depict the control signals for three different control weights: no control weight, $\zeta = 10$, and $\zeta = 5$. As expected, heavier control weights (i.e. smaller ζ) reduce the control effort. Even without any control weighting, the response in Fig. 3(a) goes to zero for large measurement deviations. The fact that this behavior is seen when there is no control weighting implies that the attenuation of the control signal for large measurement deviations is due to the cpdf and not the objective function. Moreover, this behavior is not shared by the LEG controller that uses a similar objective function but assumes light-tailed, Gaussian distributions.

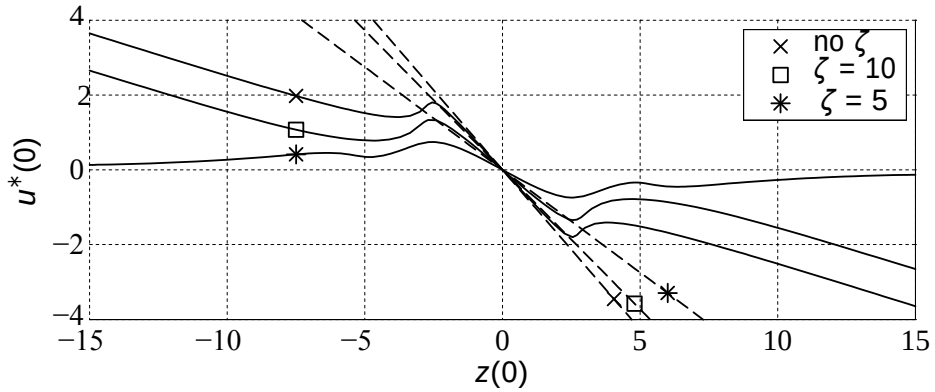
2.3.2 Multi-Step Numerical Example for the Scalar and Two-State System

The dynamic characteristics of the Cauchy optimal controller of the scalar dynamic system, obtained by maximizing the performance index in (30), are explored through several multi-step numerical examples. The Cauchy optimal predictor control results are compared against the least-squares equivalent LEG predictor controller and the Kalman filter [19, Appendix D]. The example that is discussed in this section is a stable system with $\Phi = 0.95$, $H = 1$, and a horizon length of $m = 2$. The state weight parameter is chosen as $\eta_p = \eta_{k+2} = 0.7$, while the control weights are chosen as $\zeta_i = 8$, $i = k, k+1$. The noise parameter values β and γ are interchanged to see how the controller performance changes when it is designed for a large measurement noise impulse in contrast to when it is designed for a large process noise impulse.

The simulations results are depicted in Figs. 4 where the system parameters are given. First, for $\gamma = 0.1$ and $\beta = .02$, depicted in Fig. 4(a) when the noises are small, the Cauchy and the LEG controllers exhibit similar performance. However, they behave rather differently when a large measurement pulse occurs. A measurement noise pulse does not represent a state deviation and thus, for proper regulation, the controller should ignore that measurement. The Cauchy predictive controller, designed for $\gamma > \beta$, is able to make this distinction, whereas the LEG predictive



(a) $\beta = 0.1$, $\gamma = 0.5$.

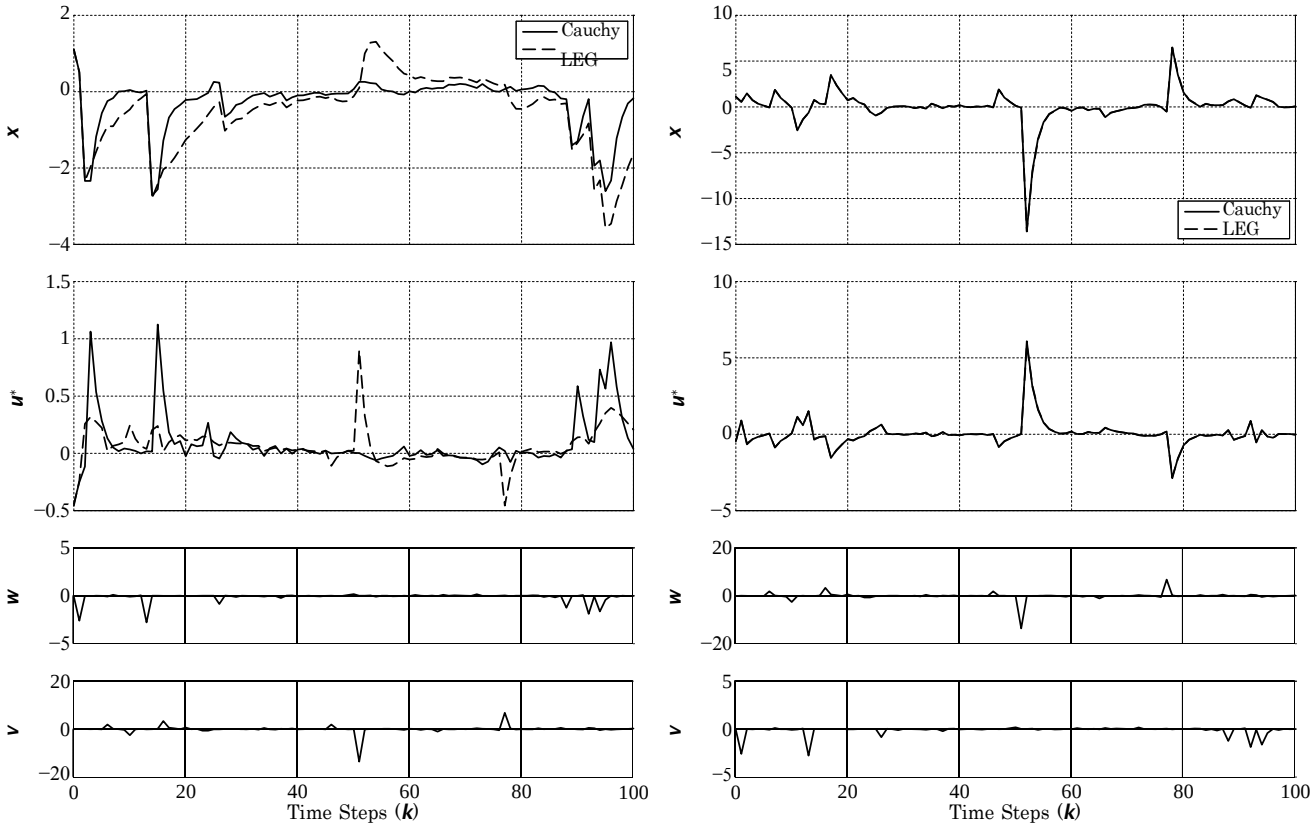


(b) $\beta = 0.5$, $\gamma = 0.1$.

Figure 3: Optimal control vs the measurement for the first time step.

controller reacts linearly to all the pulses and does not differentiate as shown in Fig. 4(a). At time steps $k = 2$ and $k = 13$ process noise pulses occur, and although both controllers react to them and are able to overcome this deviation, the Cauchy controller does so much quicker than the LEG by applying a much larger control effort. The Cauchy applies a larger control because its gain for small measurement values are higher than that of the LEG. Conversely, when a large measurement pulse occurs at $k = 51$, the Cauchy controller ignores it, applying almost zero control, whereas the LEG controller applies a very large control input that causes the state to deviate away from zero, which then required additional control effort to correct. This way the Cauchy controller manages to avoid unnecessary actuation and thus maintains the system performance. When $\gamma < \beta$, the behavior of the Cauchy and LEG controller is similar, as shown in Fig. 4(b). This demonstrates the same linear behavior as was seen in Fig. 2(b).

In [20, 21, Appendices E & F], the two-state, multi-step example uses the same dynamic system as the two-state, two-step, single measurement example where $m = 2$ and $\zeta_k = \zeta_{k+1}$. In Fig. 5(a), there is more uncertainty in the state measurement noise than in the process noise. When large measurement deviations occur (such as at $k = 52$), the Cauchy controller's effort is very small even though there is no weighting on the control inputs. In contrast, the LEG controller responds with a



(a) $H = 1,\ \alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,$
$\Phi = 0.95,\ \eta_2 = 0.7, \zeta_1 = 8$

(b) $H = 1,\ \alpha = 0.5,\ \beta = 0.1,\ \gamma = 0.02,$
$\Phi = 0.95,\ \eta_2 = 0.7, \zeta_1 = 8$

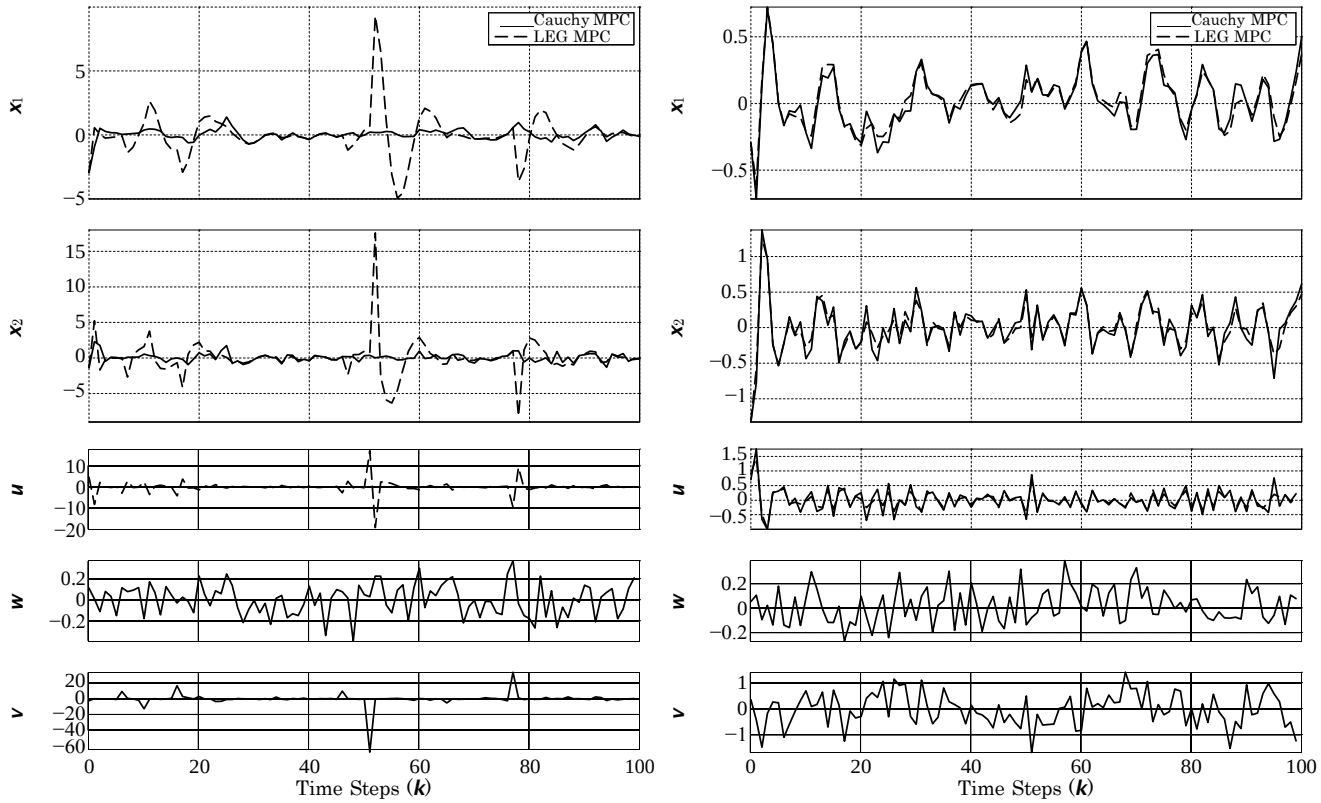
Figure 4: 2-step Cauchy and Gaussian controllers with β and γ parameters interchanged.

large control effort that drives the states from their regulated state of zero. When the measurement noise density parameter dominates the process noise density parameter in constructing the Cauchy controller, the effect of measurement outliers is mitigated, while still responding to state deviations due to process noise. In the Gaussian simulation, shown in Fig. 5(b), both controllers perform identically.

3 Conclusions

The current results of state estimation and control in linear, discrete time systems with additive Cauchy noises are summarized in References [11, 12, 14 – 18, 20, 21, 24 – 28, Appendices A-F]. Although closed-form solutions, especially for the state estimation problem, are presented in those publications, they entail a significant numerical complexity that may limit their real-time application in practical engineering systems.

The estimator complexity stems from the representation of the associated characteristic func-



(a) Cauchy Simulation (b) Gaussian Simulation

Figure 5: Cauchy and Gaussian controller performance for two-state system.

tions of ucpdf as a continuously growing sum of terms, each of which depends on the entire history of the measured data and intermediately computed parameters, as can be seen in the references and summarized briefly in Section 2.1.5. For scalar and two-state systems, compact and simplified representations were found to represent those characteristic functions [12, 16]. Moreover, for the scalar system case, efficient methods were found to prune the terms of the characteristic function while bounding the error of the approximated result [12]. However, by using a fixed data window, measurement data can be streamed continuously with insignificant error for the two-state system [17, Appendix C]. Although the determination of the controller requires a numerical optimization procedure [20, 21, Appendix E & F], the Cauchy controller demonstrates how to handle outliers, an issue that has plagued designers who use Gaussian control algorithms.

4

References

[1] N. N. Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2010.

[2] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner. Near Optimal Detection of Signals in Impulsive Noise Modeled with Asymmetric Alpha-Stable Distribution. *IEEE Communications Letters*, 2(10):282 – 284, Oct. 1998.

[3] P. Reeves. A non-Gaussian turbulence simulation. Tech. Rep. AFFDL-TR-69-67, Air Force Flight Dynamics Laboratory, 1969.

[4] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, New York, 1994.

[5] P. Tsakalides and C. L. Nikias. *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[6] G. R. Arce. *Nonlinear Signal Processing: A statistical Approach*. Wiley, New Jersey, 2005.

[7] J. P. Nolan, J. G. Gonzalez, and R.C. Nunez. Stable Filters: A robust Signal Processing Framework for Heavy-Tailed Noise. In *In Proceedings of the 2010 IEEE Radar Conference*, 2010.

[8] G. A. Tsihrintzis. *Statistical Modeling and Receiver Design for Multi-User Communication Networks;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[9] C. J. Masreliez and R. D. Martin. Robust Bayesian Estimation for the Linear Model and Robustifying the Kalman Filter. *IEEE Transactions on Automatic Control*, AC-22(3):361 – 371, June 1977.

[10] G. A. Hewer, R. D. Martin, and J. Zeh. Robust Preprocessing for Kalman Filtering of Glint Noise. *IEEE Transactions on Aerospace and Electronic Systems*, AES-23(1):120 – 128, January 1987.

[11] M. Idan and J. L. Speyer. Cauchy Estimation for Linear Scalar Systems. *IEEE Transactions on Automatic Control*, 55(6):1329 – 1342, 2010.

[12] M. Idan and J. L. Speyer. State Estimation for Linear Scalar Dynamic Systems fith Additive Cauchy Noises: Characteristic Function Approach. *SIAM Journal on Control and Optimization*, 50(4):1971 – 1994, 2012.

[13] L. N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*. John Wiley & Sons, Inc., second edition, 1994.

[14] M. Idan and J. L. Speyer. Multivariate Cauchy Estimator with Scalar Measurement and Process Noises. *SIAM Journal on Control and Optimization*, submitted.

[15] M. Idan and J. L. Speyer. Multivariate Cauchy Estimator with Scalar Measurement and Process Noises. In *IEEE CDC*, Florance, Italy, December. 2013.

[16] J. L. Speyer, M. Idan, and J. Fernández. The Two-State Estimator for Linear System with Additive Measurement and Process Cauchy Noise. In *IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2012.

[17] J. Fernández, J. L. Speyer, and M. Idan. Linear Dynamic Systems with Additive Cauchy Noises Part 1: Stochastic Estimation for Two-State Systems. *IEEE Transactions on Automatic Control*, submitted.

[18] J. L. Speyer, M. Idan, and J. Fernández. Multi-Step Prediction Optimal Control for a Scalar Linear System with Additive Cauchy Noise. In *IEEE Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

[19] J. L. Speyer, M. Idan, and J. Fernández. A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise. *Automatica*, in press.

[20] J. Fernández, J. L. Speyer, and M. Idan. A Stochastic Controller for Linear Systems with Additive Cauchy Noises. In *IEEE CDC*, Florance, Italy, December. 2013.

[21] J. Fernández, J. L. Speyer, and M. Idan. Linear Dynamic Systems with Additive Cauchy Noises Part 2: Stochastic Model Predictive Control. *IEEE Transactions on Automatic Control*, submitted.

[22] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, New York, 1975.

[23] J. L. Speyer and W. H. Chung. *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

[24] M. Idan and J. L. Speyer. Cauchy Estimation for Linear Scalar Systems. In *IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008.

[25] M. Idan and J. L. Speyer. Improvements on Cauchy Estimation for Linear Scalar Systems. In *Israel Annual Conference on Aerospace Sciences*, Tel-Aviv/Haifa, Israel, Feb. 2010.

[26] M. Idan, A. A. Emadzadeh, and J. L. Speyer. Optimal Control for a Scalar One-Step Linear System with Additive Cauchy Noise. In *American Control Conference*, Baltimore, Maryland, Jun. 2010.

[27] M. Idan and J. L. Speyer. Characteristic Function Approach for Estimation of Scalar Systems with Cauchy Noises. In *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario Canada, Aug. 2010.

[28] J. L. Speyer, M. Idan, and J. Fernández. Model Predictive Control for Scalar Linear Systems with Cauchy Noises. In *Israel Annual Conference on Aerospace Sciences*, Tel-Aviv/Haifa, Israel, Feb. 2011.

A Appendix

Moshe Idan and Jason L. Speyer
Multivariate Cauchy Estimator with Scalar Measurement and Process Noises
Submitted to SIAM Journal on Optimization and Control

MULTIVARIATE CAUCHY ESTIMATOR WITH SCALAR MEASUREMENT AND PROCESS NOISES

MOSHE IDAN* AND JASON L. SPEYER†

Abstract. The conditional mean estimator for a n-state linear system with additive Cauchy measurement and process noises is developed. It is shown that although the Cauchy densities that model the initial state, the process noise and the measurement noise have undefined first moments and an infinite second moment, the probability density function conditioned on the measurement history does have a finite conditional mean and conditional variance. For the multi-variable system state, the characteristic function of the unnormalized conditional probability density function is sequentially propagated through measurement updates and dynamic state propagation, while expressing the resulting characteristic function in a closed analytical form. Once the characteristic function of the unnormalized conditional probability density function is obtained, the probability density function of the measurement history, the conditional mean and conditional variance are easily computed from the characteristic function and its continuous first and second derivatives, evaluated at the origin in the spectral variables' domain. These closed form expressions yield the sequential state estimator. A three-state dynamic system example demonstrates numerically the performance of the Cauchy estimator.

Key words. Cauchy probability density function, non-linear estimation with heavy tailed noises, characteristic functions

AMS subject classifications. 93E11, 62L12

1. Introduction. In many engineering, economic, telecommunications, and science applications the underlying random processes or noises have significant volatility, which are not captured by Gaussian distributions [12]. Rather than light-tailed Gaussian distributions, heavy-tailed distributions have been shown to better represent these volatile random fluctuations. Examples are radar and sonar sensor noise [6] and air turbulent environment noise [9]. Our objective is to develop a filtering technique for linear dynamic systems with heavy-tailed distributed noises while using a particular case of symmetric alpha-stable ($S\alpha$-S) distributions [10]. In its simplified form, $S\alpha$-S distributions of scalar random variables are characterized by their characteristic function $\varphi(v) = e^{-\sigma^{\alpha}|v|^{\alpha}}$, where σ and α are parameters and v is the spectral variable. In this class, $\alpha = 1$ and 2 yield the Cauchy and Gaussian distributions, respectively. For $\alpha \in (0,\ 2)$, all the densities have infinite variance.

In the detection of a radar signal in clutter, the in-phase component of radar clutter time series agrees extremely well with a $S\alpha$-S probability density function (pdf) with $\alpha = 1.7$ [13]. For $\alpha \in [1, 2]$ a maximum likelihood Cauchy detector, which is in the class of myriad filters [1], exhibited performance that is very close to the Cramer-Rao bound, whereas a maximum likelihood Gaussian detector deviated significantly as α varied from 2 to 1. The myriad filters or detectors, based on a cost criterion derived from the α-stable pdfs, show significant improvement in detecting a signal in heavy tailed noise over Gaussian detectors [8], although they are numerically intensive. Techniques for enhancing the estimation robustness performance of Gaussian filters attempt to mitigate the effect of outliers due to impulsive measurement and process noise. These estimation techniques adjust to the measurement data by incorporating ad hoc heavy-tailed densities into the estimation scheme. For example, [**?**] robustifies the Kalman filter by replacing the Gaussian densities with heavy-tailed symmetric Student' s-t densities, which behave much like the Gaussian about their medians. Robust filters with data-dependent mean-square error recursions [7] have similar properties. Fortunately, our recursive estimator, without approximation, contains these robustness features and have motivated the derivation of a sequential non-linear estimator for scalar linear dynamic systems with additive Cauchy process and measurement noises [3, 4]. Although Cauchy noises do not have a well defined first moment and have infinite second and higher moments [5], the conditional pdf (cpdf) of the system state given the measurement history was determined analytically and shown to have well defined finite conditional first and second order moments [3].

Unfortunately, the recursion scheme for generating the cpdf directly for a scalar linear system [3] does not generalize for the vector state. In [4] a similar estimator was derived by generating the characteristic function of the unnormalized cpdf (ucpdf) in a recursive scheme. This approach is somewhat simpler than the scheme in [3], allows for a stronger result regarding the decay of the estimator parameters with time, and can be generalized to the multi-variable case.

The generalization of the characteristic function of the ucpdf approach [4] to the multi-variable case is the essence of this paper, which is organized as follows. We begin by formulating the estimation problem for a n-dimensional,

*Associate Professor, Faculty of Aerospace Engineering, Technion, Haifa, Israel (moshe.idan@technion.ac.il).

† Distinguished Professor, Department of Mechanical and Aerospace Engineering, University of California, Los Angeles (speyer@seas.ucla.edu).

This work was partially supported by the United States - Israel Binational Science Foundation, Grant 2008040 and by the US Air Force Office of Scientific Research, Award No. FA9550-10-1-0570.

discrete-time, linear system forced by scalar Cauchy process noise and a scalar measurement with additive Cauchy measurement noise in section 2. In section 3 we show how the characteristic function for the ucpdf of the system state conditioned on the measured history is computed sequentially for the first two measurement updates and a time propagation step. This sequential exposition suggests the general form of the characteristic function for the ucpdf given in section 4. Fundamental to obtaining this sequential estimator is the closed form solution to a convolution integral that is given in Appendix B and used for the measurement update at each time step. In section 5 we prove that under a mild condition on the system input and output matrices the characteristic function of the ucpdf is twice continually differentiable. This yields closed form analytical expressions for the minimum variance estimate of the states and the estimation error conditional variance. In section 6 a three state system is used to illustrate the performance of the sequential estimator. We offer some concluding remarks in section 7.

2. Problem Formulation. We consider the single-input-single-output multivariable linear system

$$x_{k+1} = \Phi x_k + \Gamma w_k, \qquad z_k = Hx_k + v_k, \tag{2.1}$$

with state vector $x_k \in \mathrm{R}^n$, scalar measurement z_k, and known matrices $\Phi \in \mathrm{R}^{n\times n}$, $\Gamma \in \mathrm{R}^{n\times 1}$, and $H \in \mathrm{R}^{1\times n}$. The noise inputs are assumed to be independent Cauchy distributed random variables. Specifically, w_k is assumed to be Cauchy distributed with a zero median and a scaling parameter $\beta > 0$. Similarly, v_k has a Cauchy pdf with a median of zero and a scaling parameter $\gamma > 0$. The characteristic functions of these scalar noises are assumed to be time independent and given by

$$\varphi_W(\bar{v}) = e^{-\beta|\bar{v}|}, \quad \varphi_V(\bar{v}) = e^{-\gamma|\bar{v}|}. \tag{2.2}$$

These characteristic functions have a scalar argument $\bar{v}$. The initial conditions at $k = 1$ are also assumed to be independent Cauchy distributed random variables. Specifically, each i-th element x_{1i} of the initial state vector x_1 has a Cauchy pdf with a zero median and a scaling parameter $\alpha_i > 0$, $i = 1, \ldots, n$. The characteristic function of the joint pdf of the initial conditions, which is a function of a n-dimensional spectral variable $v \in \mathrm{R}^n$, is given by

$$\varphi_{x_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i|v_i|} = \exp\left(-\sum_{i=1}^{n}\alpha_i|v_i|\right) = \exp\left(-\sum_{i=1}^{n} p_i^1|(a_i^1, v)| + j(b_1^1, v)\right). \tag{2.3}$$

The last form was introduced for notational convenience to be used in the sequel. We used the definitions

$$p_i^1 = \alpha_i, \quad a_i^1 = e_i, \quad i = 1, \ldots, n, \quad b_1^1 = \{0\}_n, \tag{2.4}$$

where e_i is a n-dimensional i-th unity vector and $\{0\}_n$ is n-dimensional vector of zeros. In fact, any distribution of the initial conditions can be handled by the derived estimator as long as its characteristic function is of the form given in (2.3) with any vectors a_i^1 and b_1^1, and parameters $p_i^1 > 0$.

The goal is to compute the minimum variance estimate of x_k given the measurement history or $y_k = [z_1 \quad z_2 \quad \cdots \quad z_k]$.

3. Initial Derivations. The method proposed to solve this Cauchy estimation problem entails propagating the characteristic function of the cpdf of the state vector given a history of measurements. Evaluating this characteristic function and its derivatives at the origin of the spectral vector v will provide the desired state estimate and its error variance. The characteristic function is initiated by the expression given in (2.3). It changes during a measurement update when a new measurement is processed, and during time propagation affected by the process noise input. We begin by showing how this characteristic function is computed for the first two measurement updates and one time propagation step. This will suggest the general form of the characteristic function and hence the general estimator.

3.1. Measurement Update at $k = 1$. Given the initial characteristic function of the system state (2.3), we are interested in computing the characteristic function of the initial state x_1 conditioned on the initial measurement $z_1 = Hx_1 + v_1$. The desired characteristic function is given by

$$\varphi_{x_1|z_1}(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{x_1|z_1}(x_1|z_1)\, e^{jv^T x_1} dx_1, \tag{3.1}$$

where the measurement history is simply $y_1 = z_1$ and the cpdf $f_{x_1|z_1}(x_1|z_1)$ is computed as

$$f_{x_1|z_1}(x_1|z_1) = \frac{f_{x_1 z_1}(x_1, z_1)}{f_{z_1}(z^1)} = \frac{f_{z_1|x_1}(z_1|x_1)\, f_{x_1}(x_1)}{f_{z_1}(z^1)} = \frac{f_V(z_1 - Hx_1)\, f_{x_1}(x_1)}{f_{z_1}(z_1)} \tag{3.2}$$

with

$$f_{Z_1}(z_1) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1Z_1}(x_1, z_1)\,dx_1 = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_V(z_1 - Hx_1)\, f_{X_1}(x_1)\,dx_1. \tag{3.3}$$

Hence,

$$\varphi_{X_1|Z_1}(v) = \frac{1}{f_{Z_1}(z_1)} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_V(z_1 - Hx_1)\, f_{X_1}(x_1)\, e^{jv^T x_1}\,dx_1. \tag{3.4}$$

The subsequent derivations can be simplified by avoiding the division by $f_{Z_1}(z_1)$ in (3.2) and (3.4). Thus we use the unnormalized cpdf and its characteristic function defined by

$$\bar{f}_{X_1|Z_1}(x_1|z_1) = f_{X_1|Z_1}(x_1|z_1)\, f_{Z_1}(z_1) = f_V(z_1 - Hx_1)\, f_{X_1}(x_1), \tag{3.5}$$

and

$$\bar{\varphi}_{X_1|Z_1}(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_V(z_1 - Hx_1)\, f_{X_1}(x_1)\, e^{jv^T x_1}dx_1. \tag{3.6}$$

When the actual (normalized) functions are needed for, e.g., computing the state estimates, the normalization factor can be easily determined by evaluating $\bar{\varphi}_{X_1|Z_1}(v)$ at $v = \{0\}_n$, i.e., $f_{Z_1}(z_1) = \bar{\varphi}_{X_1|Z_1}(\{0\}_n)$.

The integral in (3.6) resembles a Fourier transform of a product of two functions. Using the dual convolution property of Fourier transforms, it is shown in (A.7) of Appendix A that for a scalar measurement z_1, and thus a scalar function $f_V(\cdot)$, the integral in (3.6) can be expressed as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_{X_1}(v - H^T\eta)\, \varphi_V(-\eta)\, e^{jz_1\eta}d\eta, \tag{3.7}$$

where $\varphi_V(\cdot)$ and $\varphi_{X_1}(\cdot)$ were defined in (2.2) and (2.3), respectively. Using these expressions, the integral in (3.7) can be restated as

$$\begin{aligned}\bar{\varphi}_{X_1|Z_1}(v) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n} p_i^1|(a_i^1, v - H^T\eta)| + j(b_1^1, v - H^T\eta) - \gamma|\eta| + jz_1\eta\right)d\eta \\ &= \frac{e^{j(b_1, v)}}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n} p_i^1|(a_i^1, v) - Ha_i^1\eta| - \gamma|\eta| + j(z_1 - Hb_1^1)\eta\right)d\eta.\end{aligned} \tag{3.8}$$

This integral is solved using the methodology presented in Appendix B.1. For that, the coefficients of η in the absolute value term of (3.8) have to be normalized to one. Clearly, this cannot be attained if some of these coefficients are zero. Hence, first we assume that

$$Ha_i^1 \neq 0,\ i = 1, \ldots, n, \tag{3.9}$$

or equivalently that all the elements of $H = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix}$ are nonzero. With this assumption, (3.8) is restated as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{e^{j(b_1^1, v)}}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n+1} \rho_i|(\mu_i, v) - \eta_i| + j\zeta_1\eta\right)d\eta, \tag{3.10}$$

where we defined

$$\rho_i = p_i^1|Ha_i^1|, \qquad \mu_i = a_i^1/(Ha_i^1), \qquad i = 1, \ldots, n, \tag{3.11a}$$

$$\rho_{n+1} = \gamma, \qquad \mu_{n+1} = \{0\}_n, \qquad \zeta_1 = z_1 - Hb_1^1. \tag{3.11b}$$

Using the results in (B.9) and (B.10), while associating ξ_i and z in the later with (μ_i, v) and ζ_1, respectively, the integral in (3.10) is evaluated as

$$\bar{\varphi}_{X_1|Z_1}(v) = e^{j(b^1_1 - \sum_{i=1}^{n+1} g^{1|1}_i \sum_{\ell=1}^{n+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_i, v) \exp\left[\sum_{\ell=1}^{n+1} \rho_\ell |(\mu_\ell - \mu_i, v)| + j(\zeta_1 \mu_i, v)\right]} \tag{3.12}$$

$$= \sum_{i=1}^{n+1} g^{1|1}_i \, y^{1|1}_{gi}(v) \exp y^{1|1}_{ei}(v),$$

where the coefficient functions $g^{1|1}_i(\cdot)$ are given by

$$g^{1|1}_i\, y^{1|1}_{gi}(v) = \frac{1}{2\pi} \left[j\zeta_1 + \rho_i + y^{1|1}_{gi}(v)\right]^{-1} - \left[j\zeta_1 - \rho_i + y^{1|1}_{gi}(v)\right]^{-1}. \tag{3.13}$$

The arguments of $g^{1|1}_i(\cdot)$ and the exponents in (3.12) are

$$y^{1|1}_{gi}(v) = \sum_{\ell=1,\ell\neq i}^{n+1} \rho_\ell \operatorname{sgn}((\mu_\ell - \mu_i, v)) = \sum_{\ell=1,\ell\neq i}^{n+1} \rho_\ell \operatorname{sgn}((a_\ell, v)), \tag{3.14a}$$

$$y^{1|1}_{ei}(v) = -\sum_{\ell=1,\ell\neq i}^{n+1} \rho_\ell |(\mu_\ell - \mu_i, v)| + j(\zeta_1 \mu_i + b^1_1, v) = -\sum_{\ell=1,\ell\neq i}^{n+1} \rho_\ell |(a_\ell, v)| + j(b^{1|1}_i, v), \tag{3.14b}$$

where, while using (3.11), for $\ell \neq i$ we have defined

$$a_i = \mu_\ell - \mu_i = \begin{cases} \dfrac{a^1}{Ha^1} - \dfrac{a^1_i}{Ha^1} & i \neq n+1 \\ \dfrac{a^1}{Ha^1} & \ell \neq n+1 \\ -\dfrac{a^1_i}{Ha^1_i} & \end{cases}, \quad b_i = \zeta_1 \mu_i + b^{1|1}_1 = \begin{cases} a^1_i + b_1 & i \neq n+1 \\ z_1 - Hb_{11} \dfrac{a^1}{Ha_i} & \\ b_1 & i = n+1 \end{cases} \tag{3.15}$$

To simplify the notation and subsequent derivations, specifically to avoid the $\ell \neq i$ exclusion in the sums of (3.14), the elements in these sums are renumbered sequentially, i.e., $\ell \in [1, n]$ for each i-th term. The renumbered vectors a_i will be denoted by $a^{1|1}_i$. To accommodate subsequent derivations, the coefficients ρ in the sums of (3.14) will be marked differently for $y^{1|1}_{gi}(\cdot)$ and $y^{1|1}_{ei}(\cdot)$: we will use $q^{1|1}_i$ for the former and $p^{1|1}_i$ for the latter. The parameters z and ρ_i in (3.13) will be denoted by $c^{1|1}_i$ and $d^{1|1}_i$, respectively. Finally, the number of terms in the sum of (3.12) will be denoted by $n^{1|1}_t = n+1$, while the number of elements in the sums of (3.14) will be marked by $n^{1|1}_{ei} = n$. Although at this stage all the counters $n^{1|1}_{ei}$ are the same, as will be seen in the sequel, they may be different for different i-s. Therefore we have introduced i-dependent element counters $n^{1|1}_{ei}$. Moreover, although the counters $n^{1|1}_t$ and $n^{1|1}_{ei}$ seem to be related to each other (the former is $n+1$ while the latter are n), they will exhibit irregular changes and thus were introduced separately.

With these notations and renumbered parameters and vectors, $\bar{\varphi}_{X_1|Z_1}(v)$ is restated as

$$\bar{\varphi}_{X_1|Z_1}(v) = \sum_{i=1}^{n^{1|1}_t} g^{1|1}_i \, y^{1|1}_{gi}(v) \exp y^{1|1}_{ei}(v), \tag{3.16}$$

where

$$g^{1|1}_i\, y^{1|1}_{hi}(v) = \frac{1}{2\pi} \left[jc^{1|1}_i + d^{1|1}_i + y^{1|1}_{gi}(v)\right]^{-1} - \left[jc^{1|1}_{ei} - d^{1|1}_i + y^{1|1}_{hi}(v)\right]^{-1} \tag{3.17}$$

and

$i \quad gi \quad - \quad i \quad \overline{gi} \quad i$

(3.18) $$y^{1/1}_{gi}{}^{1|1}(v) = \sum_{i=1}^{n} q^{1|1}_{i} \operatorname{sgn}\ (a^{1|1}_{i}, v)\ ,\ y^{1|1}_{ei}(v) = -\ p^{1|1}_{i} |(a^{1|1}_{i}, v)| + j(b^{1|1}_{i}, v).$$

It is interesting to note that the initial characteristic function in (2.3) is also expressed in a form identical to that given in (3.16). The initial parameters are: $n_t^1 = 1$; $g_1^1 = 1$ (and thus there are no y_{g1}^1 arguments); $n_e^1 = n$; the parameters p_i^1 and the vectors a_i^1 and b_1^1 are given in (2.4).

3.1.1. Special Case: $Ha_i^1 = 0$. If one or more of the Ha_i^1-s are zero, i.e., one or more of the entries of the matrix H are zero, then the corresponding $e^{-p_i^1|(a_i^1,v)|}$ terms would come out of the integral in (3.8). However, the other manipulations used to solve this integral would remain the same. For the sake of simplicity, this is demonstrated by assuming that only $Ha_1^1 = 0$, while $Ha_i^1 \neq 0$ for $i = 2, \ldots, n$. In this case, (3.8) and (3.10) are restated as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{e^{-p_1^1|(a_1^1,v)|+j(b_1^1,v)}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\sum_{i=2}^{n+1} p_i^1|(a_i^1,v) - Ha_i^1\eta| - \gamma|\eta| + j(z_1 - Hb_1^1, \eta)\right) d\eta$$
$$= \frac{e^{-p_1^1|(a_1^1,v)|+j(b_1^1,v)}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\sum_{i=2}^{n+1} \rho_i|(\mu_i,v) - \eta| + j\zeta_1\eta\right) d\eta, \tag{3.19}$$

where the parameters $\rho_i, \mu_i, i = 2, \ldots, n$, and ζ_1 were defined in (3.11). The integral in the above is solved using the same technique as in the case when the condition in (3.9) holds, yielding the result

$$\bar{\varphi}_{X_1|Z_1}(v) = e^{-p_1^1|(a_1^1,v)|} \prod_{i=2}^{n+1} g_i^{1|1}(y_{gi}(v)) \exp(y_{ei}(v)), \tag{3.20}$$

where $g_i^{1|1}(\cdot)$ are given in (3.13). The arguments of $g_i^{1|1}(\cdot)$ and the exponents in the sum of (3.20) are

$$y_{gi}(v) = \sum_{\substack{\ell=2 \\ \ell\neq i}}^{n+1} \rho_\ell \operatorname{sgn}((a_\ell, v)) = \sum_{\ell=2}^{n} q_i^{1|1} \operatorname{sgn}(a_i^{1|1}, v), \tag{3.21a}$$

$$y_{ei}(v) = -\sum_{\substack{\ell=2 \\ \ell\neq i}}^{n+1} \rho_\ell |(a_\ell, v)| + j(b_i^{1|1}, v) = -\sum_{\ell=2}^{n} p_i^{1|1}(a_i^{1|1}, v) + j(b_i^{1|1}, v), \tag{3.21b}$$

with vectors a_i and b_i defined in (3.15). When ordering the elements sequentially, thus avoiding the $\ell \neq i$ exclusions in the sums of (3.21), we obtain the parameters $q_i^{1|1}$, $p_i^{1|1}$, and vectors $a_i^{1|1}$ and $b_i^{1|1}$.

The result in (3.20) can be cast in the form of (3.16) by the following modifications. First, the exponent $e^{-p_1^1|(a_1^1,v)|}$ outside of the sum is combined with the exponents in the sum by modifying the argument $y_{ei}(v)$, i.e., adding one more element in the sum of (3.21b). This element is denoted by $-p_1^{1|1}(a_1^{1|1}, v)$ with $p_1^{1|1} = p_1^1$ and $a_1^{1|1} = a_1^1$ for all ℓ. Consequently, the sum that defines the argument to the exponents will have n elements. Second, for consistency, the sum of (3.21a) is expanded by one by introducing $q_1^{1|1} = 0$ for all ℓ. Finally, the terms in the sum of (3.20) are ordered sequentially, starting from one. As a result, $\bar{\varphi}_{X_1|Z_1}(v)$ for this special case is expressed in a form identical to that of (3.16) when condition (3.9) holds, with the only difference being that $n_t^{1|1}$ is reduced by one to $n_t^{1|1} = n$. The manipulations above can be extended to the cases where more than one Ha_i^1 is zero, hence further reducing the number of terms used to express $\bar{\varphi}_{X_1|Z_1}(v)$.

3.2. Time Propagation to $k = 2$. Given the time propagation equation $x_2 = \Phi x_1 + \Gamma w_1$ we want to compute the characteristic function of the ucpdf of x_2 given z_1. Using the results of Appendix C, (C.4), that characteristic function is given by

$$\bar{\varphi}_{X_2|Z_1}(v) = \varphi_{X_1|Z_1}(\Phi^T v)\, \varphi_W(\Gamma^T v) = \sum_{i=1}^{n_t^{1/1}} g_i^{1|1} y_{gi}^{1|1}(\Phi^T v) \exp(y_{ei}^{1|1}(\Phi^T v))\, e^{-\beta|(\Gamma,v)|}$$
$$= \sum_{i=1}^{n_t^{1/1}} g_i^{1|1} y_{gi}^{1|1}(\Phi^T v) \exp(y_{ei}^{1|1}(\Phi^T v) - \beta|(\Gamma, v)|). \tag{3.22}$$

The last expression in (3.22) can be expressed in a form that is similar to the one given in (3.16). For that, we make the following observations that lead to the definitions of new and time propagated parameters and vectors.

1. $\bar{\varphi}_{x_2|z_1}(v)$ is expressed with the same number of terms as $\bar{\varphi}_{x_1|z_1}(v)$. Hence, we define $n_t^{2|1} = n_t^{1|1}$.
2. The coefficient functions $g_i^{1|1}(\cdot)$, or specifically the parameters $c_i^{1|1}$ and $d_i^{1|1}$, remain as in (3.17), i.e., $c_i^{2|1} = c_i^{1|1}$, $d_i^{2|1} = d_i^{1|1}$, $i \in [1, \ldots, n_t^{2|1}]$. There are only changes in the arguments of $g_i^{1|1}(\cdot)$. For notational consistency we will denote the updated coefficient functions $g_i^{2|1}(\cdot)$ while remembering that
$$g_i^{2|1}(\cdot) = g_i^{1|1}(\cdot). \tag{3.23}$$
3. The parameters $q_i^{1|1}$ and $p_i^{1|1}$ used to define $y_{gi}^{1|1}(\cdot)$ and $y_{ei}^{1|1}(\cdot)$ in (3.18) are also unchanged, i.e., $p_i^{2|1} = p_i^{1|1}$, $q_i = q_i$, $i \in [1, \ldots, n_t^{1|1}]$, $\pounds \in [1, \ldots, n_{ei}^{1|1}]$.
4. The arguments of $y_{gi}^{1|1}(\cdot)$ and $y_{ei}^{1|1}(\cdot)$ are multiplied by Φ^T, hence affecting the vectors $a_i^{1|1}$ and $b_i^{1|1}$ as
$$v_{gi}^{1|1}(\Phi^T v) = \sum_{i=1}^{n_{et}^{1|1}} q_i^{1|1} \operatorname{sgn}(a_i^{1|1}, \Phi^T v) \sum_{i=\bar{1}}^{n_{et}^{1|1}} q_i^{1|1} \operatorname{sgn}(\Phi a_i^{1|1}, v) \sum_{i=\bar{1}}^{n_{et}^{1|1}} q_i^{2|1} \operatorname{sgn}(a_i^{2|1}, v), \tag{3.24a}$$
$$v_{ei}^{1|1}(\Phi^T v) = -\sum_{i=1}^{n_{et}^{1|1}} p_i^{1|1} |(\Phi a_i^{1|1}, v)| + j(\Phi b_i^{1|1}, v) = -\sum_{i=1}^{n_{et}^{1|1}} p_i^{2|1} |(a_i^{2|1}, v)| + j(b_i^{2|1}, v). \tag{3.24b}$$
Here we used the definitions $a_i^{2|1} = \Phi a_i^{1|1}$, $b_i^{2|1} = \Phi b_i^{1|1}$, $i \in [1, \ldots, n_t^{2|1}]$, $\pounds \in [1, \ldots, n_{ei}^{1|1}]$.
5. The exponents are a function of an additional element $-\beta|(\Gamma, v)|$. Hence, the number of elements that define the new argument $y_{ei}^{2|1}(\cdot)$ increase by one, i.e., $n_{ei}^{2|1} = n_{ei}^{1|1} + 1$. The parameters and vectors that define these new elements are $p_{in_{ei}^{2|1}}^{2|1} = \beta$, $a_{in_{ei}^{2|1}}^{2|1} = \Gamma$, $i \in 1, \ldots, n_t^{2|1}]$. For consistency, and to facilitate the subsequent manipulations of the characteristic function, the number of elements in the sum of the new arguments $y_{gi}^{2|1}(\cdot)$ are also increased by one, while introducing the zero parameters $q_{in_{ei}^{2|1}}^{2|1} = 0$, $i \in [1, \ldots, n_t^{2|1}]$. Hence, with these new elements, the arguments $y_{gi}^{2|1}(\cdot)$ and $y_{ei}^{2|1}(\cdot)$ are defined as
$$v_{gi}^{2|1}(v) = \sum_{i=1}^{n_{et}^{2|1}} q_i^{2|1} \operatorname{sgn}(a_i^{2|1}, v) \tag{3.25a}$$
$$y_{ei}^{2|1}(v) = y_{ei}^{1|1}(\Phi^T v) - \beta|(\Gamma, v)| = -\sum_{i=1}^{n_{et}^{2|1}} p_i^{2|1} |(a_i^{2|1}, v)| + j(b_i^{2|1}, v). \tag{3.25b}$$

Using all the time propagated and newly defined parameters, (3.22) is restated as
$$\bar{\varphi}_{x_2|z_1}(v) = \sum_{i=1}^{n_t^{2|1}} g_i^{2|1}(y_{gi}^{2|1}(v)) \exp(y_{ei}^{2|1}(v)), \tag{3.26}$$
where
$$g_i^{2|1}(y_{gi}^{2|1}(v)) = \frac{1}{2\pi} \left(jc_i^{2|1} + d_i^{2|1} + y_{gi}^{2|1}(v)\right)^{-1} - \left(jc_i^{2|1} - d_i^{2|1} + y_{gi}^{2|1}(v)\right)^{-1} \tag{3.27}$$
with arguments $y_{gi}^{2|1}(\cdot)$ and $y_{ei}^{2|1}(\cdot)$ given in (3.25). Overall, we have obtained a form which is similar to the one in (3.16), determined after the first measurement update in the previous subsection.

Remark 3.1. *It may so happen that for a given i several vectors $a_i^{2|1}$ with $\pounds \subset [1, \ldots, n_{ei}^{2|1}]$ could be co-aligned. For example, if there are two such vectors $a_i^{2|1}$ and $a_{\bar{i}}^{2|1}$, $\pounds \neq \bar{\pounds}$, they are related by a non-zero constant θ, i.e., $a_{\bar{i}}^{2|1} = \theta a_i^{2|1}$, which implies that*
$$\operatorname{sgn}(a_{\bar{i}}^{2|1}, v) = \operatorname{sgn}(\theta) \operatorname{sgn}(a_i^{2|1}, v) \tag{3.28a}$$
$$|(a_{\bar{i}}^{2|1}, v)| = |\theta| \cdot |(a_i^{2|1}, v)|. \tag{3.28b}$$

Consequently, in the sums of (3.25) *such two terms can be combined as*

$$\sum_{i=1}^{n_{ei}^{2|1}} q_i^{2|1}\,\mathrm{sgn}\left(\langle a_i^{2|1}, v\rangle\right) + q_i^{2|1}\,\mathrm{sgn}\left(\langle a_i^{2|1}, v\rangle\right) = \bar{q}_i^{2|1}\,\mathrm{sgn}\left(\langle a_i^{2|1}, v\rangle\right) \tag{3.29a}$$

$$p_i^{2|1}\left|\langle a_i^{2|1}, v\rangle\right| + p_i^{2|1}\left|\langle a_i^{2|1}, v\rangle\right| = \bar{p}_i^{2|1}\left|\langle a_i^{2|1}, v\rangle\right| \tag{3.29b}$$

with

$$\bar{q}_i^{2|1} = q_i^{2|1} + \mathrm{sgn}(\theta)\, q_i^{2|1} \tag{3.30a}$$

$$\bar{p}_i^{2|1} = p_i^{2|1} + |\theta|\, p_i^{2|1}, \tag{3.30b}$$

thus reducing $n_{ei}^{2|1}$*, the number of elements in the sums of* (3.25)*, by one. If there are more than two co-aligned vectors* $a_i^{2|1}$*, or if there are several groups of such co-aligned vectors, the above procedure can be repeated for all those occurrences thus further reducing* $n_{ei}^{2|1}$*.*

In the sequel we will assume that all co-aligned vectors were combined. This minimizes the number of elements needed in (3.25) *and contributes to the efficiency of the resulting estimator.*

3.3. Measurement Update at $k = 2$. In this subsection we perform a second scalar measurement update at $k = 2$ using the second measurement $z_2 = Hx_2 + v_2$. Specifically, we will compute $\bar{\varphi}_{x_2|y_2}(v)$, where $y_2 = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. Using the general convolution results of (A.7) in Appendix A,

$$\bar{\varphi}_{x_2|y_2}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}_{x_2|z_1}(v - H^T\eta)\,\varphi_v(-\eta)\, e^{j z_2 \eta}\, d\eta = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}_{x_2|z_1}(v - H^T\eta)\,\exp\left(-\gamma|\eta| + j z_2 \eta\right) d\eta, \tag{3.31}$$

where from (3.26)

$$\bar{\varphi}_{x_2|z_1}(v - H^T\eta) = \sum_{i=1}^{n_t^{2/1}} g_i^{2|1}\left(y_{gi}^{2|1}(v - H^T\eta)\right)\exp\left(y_{ei}^{2|1}(v - H^T\eta)\right), \tag{3.32}$$

$g_i^{2|1}(\cdot)$ is given in (3.27), and from (3.25)

$$y_{gi}^{2|1}(v - H^T\eta) = \sum_{i=1}^{n_{et}^{2/1}} q_i^{2|1}\,\mathrm{sgn}\left(\langle a_i^{2|1}, v - H^T\eta\rangle\right), \tag{3.33a}$$

$$y_{ei}^{2|1}(v - H^T\eta) = -\sum_{i=1}^{n_{et}^{2/1}} p_i^{2|1}\left|\langle a_i^{2|1}, v - H^T\eta\rangle\right| + j\langle b_i^{2|1}, v - H^T\eta\rangle. \tag{3.33b}$$

The integral in (3.31) is solved using the methodology presented in Appendix B.2. For that, the coefficients of η in the absolute value and sign terms in (3.31), or more specifically in (3.33), has to be normalized to one, assuming those coefficients are not zero. Hence, we first assume that $Ha_i^{2|1} \neq 0,\ \forall\ i, £$. The case that several $Ha_i^{2|1} = 0$ will be addressed at the end of this section. Next, we perform the following manipulation and introduce the intermediate variables

$$\langle a_i^{2|1}, v - H^T\eta\rangle = \langle a_i^{2|1}, v\rangle - Ha_i^{2|1}\eta = Ha_i^{2|1}\left(\langle \mu_i, v\rangle - \eta\right), \tag{3.34a}$$

$$q_i^{2|1}\,\mathrm{sgn}\left(\langle a_i^{2|1}, v - H^T\eta\rangle\right) = Q_i^{2|1}\,\mathrm{sgn}\left(\langle \mu_i, v\rangle - \eta\right), \tag{3.34b}$$

$$p_i^{2|1}\left|\langle a_i^{2|1}, v - H^T\eta\rangle\right| = \rho_i^{2|1}\left|\langle \mu_i, v\rangle - \eta\right|, \tag{3.34c}$$

where

$$\mu_i = a_i^{2|1}/Ha_i^{2|1},\quad Q_i^{2|1} = q_i^{2|1}\,\mathrm{sgn}\left(Ha_i^{2|1}\right),\quad \rho_i^{2|1} = p_i^{2|1}\left|Ha_i^{2|1}\right|. \tag{3.35}$$

Using these variables, (3.33) is restated as

$$y_{gi}^{2|1}(v - H^T\eta) = \sum_{i=1}^{n_{et}^{2/1}} Q_i^{2|1}\,\mathrm{sgn}\left(\langle \mu_i, v\rangle - \eta\right), \tag{3.36a}$$

$$y_{ei}^{2|1}(v - H^T\eta) = -\sum_{i=1}^{n_{et}^{2/1}} \rho_i^{2|1}\left|\langle \mu_i, v\rangle - \eta\right| + j\left(\langle b_i^{2|1}, v\rangle - Hb_i^{2|1}\eta\right). \tag{3.36b}$$

Substituting (3.32) and (3.36) into (3.31) yields

$$\bar{\varphi}_{X_2|Y_2}(\nu) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\prod_{i=1}^{n_t^{2/1}} g_i^{2|1}\Big(\prod_{\ell=1}^{n_{et}^{2/1}} Q_i^{2|1}\operatorname{sgn}((\mu_i,\nu)-\eta)\Big) \times \exp\Big(-\sum_{i=1}^{n_{et}^{2/1}} \rho_i^{2|1}|(\mu_i,\nu)-\eta| + j(b_i^{2|1},\nu) - Hb_i^{2|1}\eta\Big)\exp(-\gamma|\eta| + j z_2\eta)\,d\eta. \tag{3.37}$$

Combining the exponents, while interchanging the integral with the summation, (3.37) is restated as

$$\bar{\varphi}_{X_2|Y_2}(\nu) = \sum_{i=1}^{n_t^{2/1}} e^{\frac{j(b_i^{2/1},\nu)}{2\pi}} \int_{-\infty}^{\infty} g_i^{2|1}\Big(\prod_{\ell=1}^{n_{ei}^{2/1}+1} Q_i^{2|1}\operatorname{sgn}((\mu_i,\nu)-\eta)\Big)\exp\Big(-\sum_{\ell=1}^{n_{ei}^{2/1}+1}\rho_i^{2|1}|(\mu_i,\nu)-\eta| + j\zeta_i\eta\Big)\,d\eta. \tag{3.38}$$

In the above we defined the new parameters

$$\rho_i^{2|1} = \gamma,\ \mu_i = \{0\}_n,\ \zeta_i = z_2 - Hb_i^{2|1},\ \ell = n_{ei}^{2|1} + 1. \tag{3.39}$$

In addition, for consistency, the number of the elements in the argument of the functions $g_i^{2|1}(\cdot)$ was increased by one by introducing $Q_i^{2|1} = 0, \ell = n_{ei}^{2|1} + 1$.

Each integral in (3.38) has the form of the general integral in (B.11), when associating ξ and z with (μ_i,ν) and ζ_i, respectively. Their solution was derived in (B.15) of Appendix B and is given by

$$\bar{\varphi}_{X_2|Y_2}(\nu) = \sum_{i=1}^{n_t^{2/1}} e^{j(b_i^{2/1},\nu)}\prod_{m=1}^{n_{ei}^{2/1}+1} g_{im}^{2|2}\Big(\prod_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} Q_i^{2|1}\operatorname{sgn}(\mu_i-\mu_{im},\nu), \sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1}\rho_i^{2|1}\operatorname{sgn}(\mu_i-\mu_{im},\nu)\Big) \times \exp\Big(-\sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1}\rho_i^{2|1}|(\mu_i-\mu_{im},\nu)| + j\zeta_i(\mu_{im},\nu)\Big) = \sum_{i=1}^{n_t^{2/1}}\prod_{m=1}^{n_{ei}^{2/1}+1} g_{im}^{2|2}\big(y_{gim1}^{2|2}(\nu), y_{gim2}^{2|2}(\nu)\big)\exp\big(y_{eim}^{2|2}(\nu)\big). \tag{3.40}$$

The coefficient functions $g_{im}^{2|2}(\cdot,\cdot)$ that have two arguments are given by

$$g_{im}^{2|2}\big(y_{gim1}^{2|2}(\nu), y_{gim2}^{2|2}(\nu)\big) = \frac{1}{2\pi}\left(\frac{g_i^{2|1}\big(y_{gim1}^{2|2}(\nu)+Q_{im}^{2|1}\big)}{j\zeta_i+\rho_{im}^{2|1}+y_{gim2}^{2|2}(\nu)} - \frac{g_i^{2|1}\big(y_{gim1}^{2|2}(\nu)-Q_{im}^{2|1}\big)}{j\zeta_i-\rho_{im}^{2|1}+y_{gim2}^{2|2}(\nu)}\right). \tag{3.41}$$

The arguments of these coefficient functions and those of the exponents in (3.44) are given by

$$v_{gim1}^{2|2}(\nu) = \sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} Q_i^{2|1}\operatorname{sgn}((\mu_i-\mu_{im},\nu)) = \sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} Q_i^{2|1}\operatorname{sgn}((a_{im},\nu)), \tag{3.42a}$$

$$v_{gim2}^{2|2}(\nu) = \sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} \rho_i^{2|1}\operatorname{sgn}((\mu_i-\mu_{im},\nu)) = \sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} \rho_i^{2|1}\operatorname{sgn}((a_{im},\nu)), \tag{3.42b}$$

$$y_{eim}^{2|2}(\nu) = -\sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} \rho_i^{2|1}|(\mu_i-\mu_{im},\nu)| + j(\zeta_i\mu_{im}+b_i^{2|1},\nu) = -\sum_{\substack{\ell=1\\ \ell\neq m}}^{n_{ei}^{2/1}+1} \rho_i^{2|1}|(a_{im},\nu)| + j(b_{im},\nu), \tag{3.42c}$$

where, while using (3.35) and (3.39), for $\ell \neq m$ we have defined

$$a_{im} = \mu_i - \mu_{im} = \begin{cases} a_i^{2|1} - \dfrac{a_{im}^{2|1}}{Ha_{im}^{2|1}} & m \neq n_{ei}^{2|1}+1,\ \ell \neq n_{ei}^{2|1}+1 \\ \dfrac{a_i^{2|1}}{Ha_i^{2|1}} & m = n_{ei}^{2|1}+1 \\ -\dfrac{a_{im}^{2|1}}{Ha_{im}^{2|1}} & \ell = n_{ei}^{2|1}+1 \end{cases} \tag{3.43a}$$

$$b_{im} = \zeta_i \mu_{im} + b_i^{2|1} = \begin{cases} z_2 - Hb_i^{2|1} \dfrac{a_{im}^{2|1}}{Ha_{im}^{2|1}} + b_i^{2|1} & m \neq n_{ei}^{2|1}+1 \\ b_i^{2|1} & m = n_{ei}^{2|1}+1 \end{cases} \tag{3.43b}$$

To avoid the double-summing over i and m in (3.40), all the $(\cdot)_{im}$ terms are now re-ordered sequentially with one index i. Its range, or the number of terms in this single sum will be $n_t^{2|2} = \sum_{i=1}^{n_t^{2|1}} (n_{ei}^{2|1}+1)$. In this reordering, one has to keep track which parent $g_i^{2|1}(\cdot)$ is used to compute the updated $g_i^{2|2}(\cdot)$. This will be done by storing the parent-term index in the variable $r_i^{2|2}$. In addition, the offsets $\pm Q_{im}^{2|1}$ used to compute $g_i^{2|2}(\cdot)$ in (3.41) will be denoted by $h_i^{2|2}$. The imaginary parameter ζ_i and the real parameter $\rho_{im}^{2|1}$ in the denominator of (3.41), when re-ordered sequentially, will be denoted by $c_i^{2|2}$ and $d_i^{2|2}$, respectively. The re-ordered arguments $y_{gim1}^{2|2}(v), y_{gim2}^{2|2}(v)$, and $y_{eim}^{2|2}(v)$ will be denoted by $y_{gi1}^{2|2}(v)$, $y_{gi2}^{2|2}(v)$, and $y_{ei}^{2|2}(v)$, respectively. These arguments are defined in (3.42) by sums that have the $\ell \neq m$ exclusion. To avoid that, the parameters in these sums will be re-ordered sequentially, while accounting also for the sequential (i,m) ordering discussed above. The number of terms in those sums will be denoted by $n_{ei}^{2|2}$. The re-ordered vectors a_{im} and $b_i^{2|1}$ will be denoted by $a_i^{2|2}$ and $b_i^{2|2}$, respectively. The re-ordered parameters $Q_i^{2|1}$ of (3.42a) and $\rho_i^{2|1}$ of (3.42b) will be denoted by $q_{i1}^{2|2}$ and $q_{i2}^{2|2}$, respectively, while $\rho_i^{2|1}$ of (3.42c) will be marked by $p_i^{2|2}$.

With these substitutions, (3.40), (3.41) and (3.42) are restated as

$$\bar{\varphi}_{X_2|Y_2}(v) = \sum_{i=1}^{n_t^{2|2}} g_i^{2|2}\left(y_{gi1}^{2|2}(v), y_{gi2}^{2|2}(v)\right) \exp\left(y_{ei}^{2|2}(v)\right), \tag{3.44}$$

$$g_i^{2|2}\left(y_{gi1}^{2|2}(v), y_{gi2}^{2|2}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{2|2}}^{2|1}\left(y_{gi1}^{2|2}(v)+h_i^{2|2}\right) - g_{r_i^{2|2}}^{2|1}\left(y_{gi1}^{2|2}(v)-h_i^{2|2}\right)}{jc_i^{2|2}} + d_i^{2|2} + y_{gi2}^{2|2}(v)\right], \tag{3.45}$$

and

$$y_{gi1}^{2|2}(v) = \sum_{\ell=1}^{n_{et}^{2|2}} q_{i1}^{2|2}\,\mathrm{sgn}\left(a_\ell^{2|2}, v\right), \quad y_{gi2}^{2|2}(v) = \sum_{\ell=1}^{n_{et}^{2|2}} q_{i2}^{2|2}\,\mathrm{sgn}\left(a_\ell^{2|2}, v\right), \tag{3.46a}$$

$$y_{ei}^{2|2}(v) = -\sum_{\ell=1}^{n_{et}^{2|2}} p_i^{2|2}\left(a_\ell^{2|2}, v\right) + j\left(b_\ell^{2|2}, v\right). \tag{3.46b}$$

Note that the coefficient functions $g_{im}^{2|2}(\cdot,\cdot)$ have two arguments. It is grouped into one vector argument with two components denoted by $y_{gi}^{2|2}(v) = \begin{bmatrix} y_{gi1}^{2|2}(v) & y_{gi2}^{2|2}(v) \end{bmatrix}^T$. It is defined by combining the two equations of (3.46a) while using a two dimensional vector of parameters $q_i^{2|2} = \begin{bmatrix} q_{i1}^{2|2} & q_{i2}^{2|2} \end{bmatrix}^T$. With these definitions, (3.44) and (3.45) are restated as

$$\bar{\varphi}_{X_2|Y_2}(v) = \sum_{i=1}^{n_t^{2|2}} g_i^{2|2}\left(y_{gi}^{2|2}(v)\right) \exp\left(y_{ei}^{2|2}(v)\right), \tag{3.47}$$

$$g_i^{2|2}\left(y_{gi}^{2|2}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{2/2}}^{2|1}\left(y_{gi1}^{2|2}(v)+h_i^{2|2}\right)}{jc_i^{2|2}+d_i^{2|2}+y_{gi2}^{2|2}(v)} - \frac{g_{r_i^{2/2}}^{2|1}\left(y_{gi1}^{2|2}(v)-h_i^{2|2}\right)}{jc_i^{2|2}-d_i^{2|2}+y_{gi2}^{2|2}(v)}\right], \tag{3.48}$$

where $y_{git}^{2|2}(\cdot)$, $t = 1, 2$, denotes the t-th element of the vector $y_{gi}^{2|2}(\cdot)$. Similarly, (3.46) are restated as

$$y_{gi}^{2|2}(v) = \sum_{\ell=1}^{n_{ei}^{2|2}} q_i^{2|2}\,\mathrm{sgn}\left(a_i^{2|2}, v\right) \in \mathrm{R}^2, \quad y_{ei}^{2|2}(v) = -\sum_{\ell=1}^{n_{ei}^{2|2}} p_i^{2|2}\left|\left(a_i^{2|2}, v\right)\right| + j\left(b_i^{2|2}, v\right). \tag{3.49}$$

Remark 3.2. *Note that the coefficient functions $g_i^{2|2}(\cdot)$ and thus their parent functions $g_{r_i^{2/2}}^{2|1}(\cdot)$ are computed using the updated arguments $y_{gi}^{2|2}(v)$ of* (3.49)*. Hence, especially when evaluating the parent $g_{r_i^{2/2}}^{2|1}(\cdot)$, one does not use the old arguments $y_{gi}^{2|1}(v)$ of* (3.25)*, but rather the updated ones of* (3.49) *with the offsets $h_i^{2|2}$.*

Remark 3.3. *Due to* (3.23) *one could restate* (3.48) *as*

$$g_i^{2|2}\left(y_{gi}^{2|2}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{2/2}}^{1|1}\left(y_{gi1}^{2|2}(v)+h_i^{2|2}\right)}{jc_i^{2|2}+d_i^{2|2}+y_{gi2}^{2|2}(v)} - \frac{g_{r_i^{2/2}}^{1|1}\left(y_{gi1}^{2|2}(v)-h_i^{2|2}\right)}{jc_i^{2|2}-d_i^{2|2}+y_{gi2}^{2|2}(v)}\right], \tag{3.50}$$

which implies that one would have to store only the parameters $c_{r_i^{2/2}}^{1|1}$ and $d_{r_i^{2/2}}^{1|1}$ of the measurement updated coefficient function $g_{r_i^{2/2}}^{1|1}(\cdot)$ and not the time propagated ones (which are in fact the same.)

Remark 3.4. *Similar to the discussion presented in Remark 3.1, for a given index i, several vectors $a_i^{2|2}$ with $\ell \subset [1, \ldots, n_{ei}]$ could be co-aligned. In such cases, in the sums of* (3.49) *the elements with co-aligned $a_i^{2|2}$-s can be combined using the steps presented in Remark 3.1, thus reducing the number of elements in those sums, i.e., reducing $n_{ei}^{2|2}$. For numerical efficiency, we assume that all those co-aligned directions are combined, hence minimizing the number of elements $n_{ei}^{2|2}$ in the sums of* (3.49)*.*

Remark 3.5. *The procedure for handling the special case when one or more of the $Ha_i^{2|1}$-s are zero is similar to what was presented in subsection 3.1.1. The details are not presented here for brevity.*

In summary, the measurement updated $\bar{\varphi}_{X_2|Y_2}(v)$ of (3.44) is expressed by $n_{t}^{2|2}$ terms. The coefficient functions $g_i^{2|2}(\cdot)$ are defined in (3.50) by the set of parameters $c_i^{2|2}$ and $d_i^{2|2}$, the offsets $h_i^{2|2}$, and the index $r_i^{2|1}$ of the parent-term. Alternatively, the index $r_i^{2|2}$ can be replaced by storing the parameters that define $g_{r_i^{2/2}}^{1|1}(\cdot)$, i.e., $c_{r_i^{2/2}}^{1|1}$ and $d_{r_i^{2/2}}^{1|1}$. The two dimensional vector input arguments of $g_i^{2|2}(\cdot)$ and the scalar arguments of the exponentials in (3.44) are determined by (3.49) while using the measurement updated vectors $a_i^{2|2}$ and $b_i^{2|2}$, together with the vector parameters $q_i^{2|2}$ and scalar parameters $p_i^{2|2}$.

3.4. Summary of Initial Results. The above derivations demonstrate that the characteristic function of the ucpdf of the state x_k at time steps $k = 1$ and 2 is expressed as a sum of n_t weighted exponential terms. The number of terms increases during a measurement update and are unchanged during a time propagation step. In each such term i, the exponents and their weights or coefficients are functions of a sum of n_{ei} elements. The number n_{ei} of elements normally increases during the time propagation step and are unchanged during a measurement update. (In special cases, the number of elements may be reduced - see Remarks 3.1 and 3.4.) Moreover, it is observed that the scalar coefficients of the exponents are functions of a vector of parameters. The dimension of these vectors are unchanged during the time propagation step, and increase by one each time that a new measurement is processed. Those observations provide the insight and guidance on how to construct the characteristic functions at any time step k, as is discussed next.

4. Time Propagation and Measurement Update: General Case. The initial results presented in the previous section suggest the general form of the characteristic function of the ucpdf of the state at any time step k given past measurements up to time $k-1$ and k. Specifically, based on the measurement updated characteristic functions obtained in (3.16) and (3.44) for $k = 1$ and 2, respectively, we assume now that at any time step the characteristic function of the ucpdf of the state x_k given all the data history up to that time, i.e., $y_k =$

$z_1 \quad z_2 \quad \cdots \quad z_k$, is expressed as

$$\bar{\varphi}_{X_k|Y_k}(v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(v)\right) \exp\left(y_{ei}^{k|k}(v)\right), \tag{4.1}$$

where

$$g_i^{k|k}\left(y_{gi}^{k|k}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{k|k}}^{k-1|k-1}\left(y_{gi1}^{k|k}(v)+h_i^{k|k}\right)}{jc_i^{k|k}+d_i^{k|k}+y_{gi2}^{k|k}(v)} - \frac{g_{r_i^{k|k}}^{k-1|k-1}\left(y_{gi1}^{k|k}(v)-h_i^{k|k}\right)}{jc_i^{k|k}-d_i^{k|k}+y_{gi2}^{k|k}(v)}\right], \tag{4.2}$$

and

$$y_{gi}^{k|k}(v) = \sum_{i=1}^{n_{ei}^{k|k}} q_i^{k|k}\,\mathrm{sgn}\left(\langle a_i^{k|k}, v\rangle\right) \in \mathbb{R}^k, \quad y_{ei}^{k|k}(v) = -\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k}\left|\langle a_i^{k|k}, v\rangle\right| + j\langle b_i^{k|k}, v\rangle. \tag{4.3}$$

Here, $y_{gi}^{k|k}(\cdot)$ and $q_i^{k|k}$ are k dimensional vectors. When evaluating $g_i^{k|k}\left(y_{gi}^{k|k}(\cdot)\right)$ in (4.2), the argument $y_{gi}^{k|k}(\cdot)$ is partitioned as follows: $y_{gi1}^{k|k}(\cdot)$ is a $k-1$ dimensional vector constructed from the first $k-1$ components of $y_{gi}^{k|k}(\cdot)$, while the scalar $y_{gi2}^{k|k}(\cdot)$ is the last component of $y_{gi}^{k|k}(\cdot)$. Based on the results in subsection 3.1, at $k=1$ the numerators in (4.2) are one. Alternatively, the form of (4.2) can be maintained also for $k=1$ by initializing $g_i^{k|k}(\cdot)$ as $g_i^{0|0}(v)=1$.

Now we will perform one time propagation and one measurement update and show that the above form is maintained at any time step. Although the derivations are similar to those presented in the previous section, a full derivation is presented here for clarity.

4.1. Time Propagation from k to $k+1$. Starting with the characteristic function $\bar{\varphi}_{X_k|Y_k}(v)$ of (4.1) we want to compute $\bar{\varphi}_{X_{k+1}|Y_k}(v)$ while accounting for the time propagation equation $x_{k+1}=\Phi x_k+\Gamma w_k$. The result in (C.4) of Appendix C indicates that the desired characteristic function is given by

$$\begin{aligned}\bar{\varphi}_{X_{k+1}|Y_k}(v) &= \bar{\varphi}_{X_k|Y_k}(\Phi^T v)\,\varphi_w(\Gamma^T v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\Phi^T v)\right)\exp\left(y_{ei}^{k|k}(\Phi^T v)\right)e^{-\beta|\langle\Gamma,v\rangle|} \\ &= \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\Phi^T v)\right)\exp\left(y_{ei}^{k|k}(\Phi^T v)-\beta|\langle\Gamma,v\rangle|\right).\end{aligned} \tag{4.4}$$

The coefficient functions $g_i^{k|k}(\cdot)$ are given in (4.2). Their arguments and those of the exponents in (4.4) can be redefined as follows

$$y_{gi}^{k+1|k}(v) = y_{gi}^{k|k}(\Phi^T v) = \sum_{i=1}^{n_{ei}^{k|k}} q_i^{k|k}\,\mathrm{sgn}\left(\langle\Phi a_i^{k|k}, v\rangle\right) \in \mathbb{R}^k, \tag{4.5a}$$

$$y_{ei}^{k+1|k}(v) = y_{ei}^{k|k}(\Phi^T v)-\beta|\langle\Gamma,v\rangle| = -\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k}\left|\langle\Phi a_i^{k|k}, v\rangle\right| - \beta|\langle\Gamma,v\rangle| + j\langle b_i^{k|k}, v\rangle. \tag{4.5b}$$

With these definitions, (4.4) can be restated as

$$\bar{\varphi}_{X_{k+1}|Y_k}(v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k+1|k}(v)\right)\exp\left(y_{ei}^{k+1|k}(v)\right), \tag{4.6}$$

which clearly has a similar form as (4.1). Specifically, defining the time propagated parameters similarly to the five-steps procedure outlined in section 3.2 including the new parameters $r_i^{k+1|k}=r_i^{k|k}$, $h_i^{k+1|k}=h_i^{k|k}$, the result above can be restated as

$$\bar{\varphi}_{X_{k+1}|Z_k}(v) = \sum_{i=1}^{n_t^{k+1|k}} g_i^{k+1|k}\left(y_{gi}^{k+1|k}(v)\right)\exp\left(y_{ei}^{k+1|k}(v)\right), \tag{4.7}$$

where

$$g_i^{k+1|k}\left(y_{gi}^{k+1|k}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{k+1/k}}^{k-1|k-1}\left(y_{gi1}^{k+1|k}(v)+h_i^{k+1|k}\right)}{jc_i^{k+1|k}+d_i^{k+1|k}+y_{gi2}^{k+1|k}(v)} - \frac{g_{r_i^{k+1/k}}^{k-1|k-1}\left(y_{gi1}^{k+1|k}(v)-h_i^{k+1|k}\right)}{jc_i^{k+1|k}-d_i^{k+1|k}+y_{gi2}^{k+1|k}(v)}\right] \tag{4.8}$$

and

$$y_{gi}^{k+1|k}(v) = \sum_{=1}^{n_{ei}^{k+1/k}} q_i^{k+1|k}\operatorname{sgn}\left(a_i^{k+1|k}, v\right) \in \mathrm{R}^k, \quad y_{ei}^{k+1|k}(v) = -\sum_{=1}^{n_{ei}^{k+1/k}} p_i^{k+1|k}\left|\left(a_i^{k+1|k}, v\right)\right| + j\left(b_i^{k+1|k}, v\right). \tag{4.9}$$

The vector $y_{gi}^{k+1|k}(\cdot)$ is partitioned into two parts used to evaluate (4.8): $y_{k|k}^{gi1}(\cdot)$ is constructed from the first $k-1$ components of $y_{gi}^{k|k}(\cdot)$, while the scalar $y_{gi2}^{k}(\cdot)$ is the last component of $y_{gi}^{k|k}(\cdot)$.

The number of terms that define $\bar{\varphi}_{x_{k+1}|z_k}(\cdot)$ in (4.7) is identical to that of $\bar{\varphi}_{x_k|z_k}(\cdot)$ in (4.1). However, the number of elements needed to define the arguments of those terms has increased by one, i.e., $n_{ei}^{k+1|k} = n_{ei}^{k|k} + 1$. It may happen that for a given i, several vectors $a_i^{k+1|k}$ with $\mathcal{L} \subset [1, \ldots, n_{ei}^{k+1|k}]$ in (4.9) could be co-aligned. In such cases, the associated elements in the sums of (4.9) could be combined, thus reducing the number of elements in those sums, as discussed in Remark 3.1.

4.2. Measurement Update at $k+1$. After performing the general time propagation step in the previous subsection, thus attaining the characteristic function $\bar{\varphi}_{x_{k+1}|z_k}(\cdot)$ of (4.7), the next step is to use the next measurement update $z_{k+1} = Hx_{k+1} + v_{k+1}$ to determine $\varphi_{x_{k+1}|Y_{k+1}}(\cdot)$, where $y_{k+1} = \begin{bmatrix} z_1 & z_2 & \cdots & z_{k+1} \end{bmatrix}$. Using the general convolution results of (A.7) in Appendix A,

$$\begin{aligned}\bar{\varphi}_{x_{k+1}|Y_{k+1}}(v) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}_{x_{k+1}|Y_k}\left(v - H^T\eta\right)\varphi_v(-\eta)\, e^{jz_{k+1}\eta}\, d\eta \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}_{x_{k+1}|Y_k}\left(v - H^T\eta\right)\exp\left(-\gamma|\eta| + jz_{k+1}\eta\right) d\eta,\end{aligned} \tag{4.10}$$

where from (4.7)

$$\bar{\varphi}_{x_{k+1}|Y_k}\left(v - H^T\eta\right) = \sum_{i=1}^{n_t^{k+1/k}} g_i^{k+1|k}\left(y_{gi}^{k+1|k}\left(v - H^T\eta\right)\right)\exp\left(y_{ei}^{k+1|k}\left(v - H^T\eta\right)\right), \tag{4.11}$$

$g_i^{k+1|k}(\cdot)$ is given in (4.8), and

$$y_{gi}^{k+1|k}\left(v - H^T\eta\right) = \sum_{=1}^{n_{ei}^{k+1/k}} q_i^{k+1|k}\operatorname{sgn}\left(a_i^{k+1|k}, v - H^T\eta\right) \in \mathrm{R}^k, \tag{4.12a}$$

$$y_{ei}^{k+1|k}\left(v - H^T\eta\right) = -\sum_{=1}^{n_{ei}^{k+1/k}} p_i^{k+1|k}\left|\left(a_i^{k+1|k}, v - H^T\eta\right)\right| + j\left(b_i^{k+1|k}, v - H^T\eta\right). \tag{4.12b}$$

The integral in (4.10) is solved using the methodology presented in Appendix B.2 in precisely the same way that (3.31) was solved in section 3.3 while replacing $2|1$ by $k+1|k$ and $2|2$ by $k+1|k+1$. The equivalent forms to (3.47), (3.48) and (3.49) are

$$\bar{\varphi}_{x_{k+1}|Y_{k+1}}(v) = \sum_{i=1}^{n_t^{k+1/k+1}} g_i^{k+1|k+1}\left(y_{gi}^{k+1|k+1}(v)\right)\exp\left(y_{ei}^{k+1|k+1}(v)\right), \tag{4.13}$$

$$g_i^{k+1|k+1}\left(y_{gi}^{k+1|k+1}(v)\right) = \frac{1}{2\pi}\left[\frac{g_{r_i^{k+1/k+1}}^{k|k}\left(y_{gi1}^{k+1|k+1}(v)+h^{k+1|k+1}\right)}{jc_i^{k+1|k+1}+d_i^{k+1|k+1}+y_{gi2}^{k+1|k+1}(v)} - \frac{g_{r_i^{k+1/k+1}}^{k|k}\left(y_{gi1}^{k+1|k+1}(v)-h_i^{k+1|k+1}\right)}{jc_i^{k+1|k+1}-d_i^{k+1|k+1}+y_{gi2}^{k+1|k+1}(v)}\right], \tag{4.14}$$

(4.15a) $$y_{ai}^{k+1|k+1}(v) = \sum_{=1}^{n_{ei}^{k+1/k+1}} q_i^{k+1|k+1} \operatorname{sgn}\left(a_i^{k+1|k+1}, v\right) \in R^{k+1},$$

(4.15b) $$y_{ei}^{k+1|k+1}(v) = -\sum_{=1}^{n_{ei}^{k+1/k+1}} p_i^{k+1|k+1} \left|\left(a_i^{k+1|k+1}, v\right)\right| + j\left(b_i^{k+1|k+1}, v\right).$$

As discussed earlier in this work, if for a given index i several vectors $a_i^{k+1|k+1}$, $\pounds \subset [1, \ldots, n_{ei}^{k+1|k+1}]$ are co-aligned, the associated elements in the sums of (4.15) could be combined, consequently reducing $n_{ei}^{k+1|k+1}$, the number of elements in those sums.

The above results clearly show that the form of the characteristic function proposed in (4.1-4.3) for the time step k is also maintained at $k+1$.

5. Conditional Mean and Estimation Error Variance. The minimum conditional variance estimator of x_k given the measurement sequence $y_k = z_1\, z_2 \cdots z_k$ is the conditional mean of x_k given y_k. It can be determined by evaluating the characteristic function of (4.1) and its derivatives at $v = \{0\}_n$, or as $v \to \{0\}_n$. In this section we show that $\bar{\varphi}_{X_k|Y_k}(\cdot)$ is twice continuously differentiable and give explicit expressions for the conditional mean and the estimation error variance.

5.1. Continuity of the First Two Derivatives of the Characteristic Function. The continuity of the first two derivatives of $\bar{\varphi}_{X_k|Y_k}(\cdot)$ is proven by induction. First the case that $h_i \neq 0$, $i = \{1, \ldots, n\}$ is addressed. The case when $h_i = 0$ for some $i \subset \{1, \ldots, n\}$ is addressed at the end of this subsection.

5.1.1. Continuity of the First Two Derivatives of $\bar{\varphi}_{X_1|Y_1}$. Assuming that $h_i \neq 0$ for $i = \{1, \ldots, n\}$, the characteristic function for the ucpdf of initial state given the first measurement is given by the convolution integral in (3.10). Then, using the definitions in (2.4) and (3.11), (3.10) is rewritten as

(5.1) $$\bar{\varphi}_{X_1|Z_1}(\tilde{v}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \prod_{i=1}^{n} e^{-\alpha_i|\tilde{v}_i-\eta|} e^{-\gamma|\eta|+jz_1\eta} d\eta$$

where $\tilde{v}_i = (\mu_i, v)$ and $\tilde{v} = \begin{bmatrix} \tilde{v}_1 & \cdots & \tilde{v}_n \end{bmatrix}^T$. The integrand of (5.1) is a continuous function of η and $\tilde{v}_i \forall i = 1, \ldots, n$. Moreover, its partial derivative with respect to any $\tilde{v}_j$, given by

(5.2) $$\frac{\partial}{\partial \tilde{v}_j}\left(\prod_{i=1}^{n} e^{-\rho_i|\tilde{v}_i-\eta|} e^{-\gamma|\eta|+jz_1\eta}\right) = -\rho_j \operatorname{sgn}(\tilde{v}_j-\eta) e^{-\rho_j|\tilde{v}_j-\eta|} \prod_{\substack{i=1 \\ i\neq j}}^{n} e^{-\rho_i|\tilde{v}_i-\eta|} e^{-\gamma|\eta|+jz_1\eta},$$

is piecewise continuous, bounded, and integrable. Hence, when computing the partial derivative of $\bar{\varphi}_{X_1|Z_1}(\tilde{v})$ with respect to $\tilde{v}_j$, the differentiation and integration operations can be reversed [2] to yield

(5.3) $$\frac{\partial \bar{\varphi}_{X_1|Z_1}(\tilde{v})}{\partial \tilde{v}_j} = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \rho_j \operatorname{sgn}(\tilde{v}_j-\eta) e^{-\rho_j|\tilde{v}_j-\eta|} \prod_{\substack{i=1 \\ i\neq j}}^{n} e^{-\rho_i|\tilde{v}_i-\eta|} e^{-\gamma|\eta|+jz_1\eta} d\eta$$

Since the integrand in (5.3) is piecewise continuous and bounded functions of η and $\tilde{v}_i \forall i = 1, \ldots, n$, the first derivative $\partial \bar{\varphi}_{X_1|Z_1}(\tilde{v})/\partial \tilde{v}$ is continuous [2].

The second partial derivatives of $\bar{\varphi}_{X_1|Z_1}(\tilde{v})$ is attained by differentiating (5.3). To avoid the differentiation of the piecewise continuous function $\operatorname{sgn}(\tilde{v}_j - \eta)$, we introduce the change of variables $\tilde{v}_j - \eta = \sigma \Rightarrow \eta = \tilde{v}_j - \sigma,\ d\eta = -d\sigma$ to restate (5.3) as

(5.4) $$\frac{\partial \bar{\varphi}_{X_1|Z_1}(\tilde{v})}{\partial \tilde{v}_j} = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \rho_j \operatorname{sgn}(\sigma) e^{-\rho_j|\sigma|} \prod_{\substack{i=1 \\ i\neq j}}^{n} e^{-\rho_i|\tilde{v}_i-\tilde{v}_j+\sigma|} e^{-\gamma|\tilde{v}_j-\sigma|+jz_1(\tilde{v}_j-\sigma)} d\sigma.$$

The integrand in (5.4) is a piecewise continuous in σ, continuous in $\tilde{v}_i \forall i = 1, \ldots, n$, and its first partial derivative with respect to $\tilde{v}_j$ is piecewise continuous, bounded, and integrable. Consequently, in computing the partial derivative

of (5.4) with respect to $\tilde{v}$, the order of differentiation and integration can be interchanged as

$$\frac{\partial^2 \bar{\varphi}_{X_1|Z_1}(\tilde{v})}{\partial \tilde{v} \partial \tilde{v}_j} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \rho_j \operatorname{sgn}(\sigma) e^{-\rho_j|\sigma|} \frac{\partial}{\partial \tilde{v}} \left[\prod_{\substack{i=1 \\ i \neq j}}^{n} e^{-\rho_i|\tilde{v}_i - \tilde{v}_j + \sigma|} e^{-\gamma|\tilde{v}_j - \sigma| + jz_1(\tilde{v}_j - \sigma)} \right] d\sigma. \tag{5.5}$$

The partial derivative and hence the integrand in (5.5) is a piecewise continuous and bounded function of the variables σ and $\tilde{v}_i \forall i = 1, \ldots, n$. Therefore, the integrals exist for all j and $\pounds$ and are continuous with respect to all $(\tilde{v}_j, \tilde{v})$ [2]. This implies that all the elements of the Hessian of $\bar{\varphi}_{X_1|Z_1}(\tilde{v})$ are continuous for all $\tilde{v}$.

In summary, $\bar{\varphi}_{X_1|Z_1}(\tilde{v})$ is twice continuously differentiable with respect to $\tilde{v}$. Since v is a linear function of $\tilde{v}$, $\bar{\varphi}_{X_1|Z_1}(v)$ is twice continuously differentiable with respect to v.

5.1.2. Continuity of the First Two Derivatives of $\bar{\varphi}_{X_2|Y_2}$. The time propagated characteristic function at $k = 2$, which was given in (3.22), is considered first. It is rewritten as

$$\bar{\varphi}_{X_2|Z_1}(v) = \bar{\varphi}_{X_1|Z_1}(\Phi^T v)\, e^{-\beta|\Gamma^T v|}. \tag{5.6}$$

Since, as it was shown above, $\bar{\varphi}_{X_1|Z_1}(v)$ is twice continuously differentiable, $\bar{\varphi}_{X_1|Z_1}(\Phi^T v)$ is also twice continuously differentiable for any transition matrix Φ. Clearly, $\bar{\varphi}_{X_2|Z_1}(v)$ is continuous, being a product of two continuous functions. However, since the first derivative of $e^{-\beta|\Gamma^T v|}$ is not continuous at $v = \{0\}_n$, the associated pdf $f_{X_2|Z_1}(x_2|z_1)$ does not have any moments. This implies that we cannot compute *a priori* estimates of the state x_2 given only the past measurement z_1.

The characteristic function at the second measurement update, given the measurement history $y_2 = \{z_1, z_2\}$, was expressed in (3.31) and rewritten here using the explicit form of (5.6) as

$$\bar{\varphi}_{X_2|Y_2}(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\varphi}_{X_1|Z_1}(v - H^T \eta) e^{-\beta|\Gamma^T v - H\Gamma\eta|} e^{-\gamma|\eta| + jz_2\eta}\, d\eta, \tag{5.7}$$

where we used the notation $\tilde{\varphi}_{X_1|Z_1}(v) = \bar{\varphi}_{X_1|Z_1}(\Phi^T v)$. Note that if $H\Gamma = 0$, the term $e^{-\beta|\Gamma^T v|}$ would come out of the integral in (5.7). Consequently, in this case $\bar{\varphi}_{X_2|Z_2}(v)$ would not be continuously differentiable with respect to v, and there would be no minimum variance estimate of x_k given y_k for all $k \geq 2$. Therefore, $H\Gamma \neq 0$ is a necessary condition for the continuous differentiability of $\bar{\varphi}_{X_2|Y_2}(v)$ and the existence of the desired estimate of the state x_k at any time step k.

From section 5.1.1, $\tilde{\varphi}_{X_1|Z_1}(v - H^T \eta)$ is twice continuously differentiable with respect to v. Moreover, $e^{-\beta|\Gamma^T v - H\Gamma\eta|}$ is once piecewise continuously differentiable with respect to v as

$$\frac{\partial e^{-\beta|\Gamma^T v - H\Gamma\eta|}}{\partial v} = -\beta \operatorname{sgn}(\Gamma^T v - H\Gamma\eta)\, e^{-\beta|\Gamma^T v - H\Gamma\eta|} \Gamma^T. \tag{5.8}$$

Therefore, when constructing the first derivative of $\bar{\varphi}_{X_2|Y_2}(v)$ with respect to v, the order of differentiation and integration in can reversed to yield

$$\begin{aligned} \frac{\partial \bar{\varphi}_{X_2|Y_2}(v)}{\partial v} = {} & \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial \tilde{\varphi}_{X_1|Z_1}(v - H^T \eta)}{\partial v} e^{-\beta|\Gamma^T v - H\Gamma\eta|} e^{-\gamma|\eta| + jz_2\eta}\, d\eta \\ & + \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\varphi}_{X_1|Z_1}(v - H^T \eta) \left[-\beta \operatorname{sgn}(\Gamma^T v - H\Gamma\eta)\, e^{-\beta|\Gamma^T v - H\Gamma\eta|} \Gamma^T \right] e^{-\gamma|\eta| + jz_2\eta}\, d\eta. \end{aligned} \tag{5.9}$$

Since the integrands of (5.9) are piecewise continuous and bounded functions of η and $v_i, i = 1, \ldots, n$, $\partial \bar{\varphi}_{X_2|Y_2}(v)/\partial v$ is continuous.

The second partial derivative of $\bar{\varphi}_{X_2|Y_2}(v)$ can be taken after making the following change of variables

$$\frac{\Gamma^T v}{H\Gamma} - \eta = \sigma \Rightarrow \eta = \frac{\Gamma^T v}{H\Gamma} - \sigma,\ d\eta = -d\sigma. \tag{5.10}$$

Substituting (5.10) into (5.9) we obtain

$$\frac{\partial \bar{\varphi}_{X_2|Y_2}(v)}{\partial v} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\partial \tilde{\varphi}_{X_1|Z_1}(Pv + H^T\sigma)}{\partial v} e^{-\beta|H\Gamma\sigma|} e^{-\gamma\sqrt{\frac{\Gamma^T v}{H\Gamma} - \sigma + jz_2}\left(\frac{\Gamma^T v}{H\Gamma} - \sigma\right)}\, d\sigma$$
$$- \frac{1}{2\pi}\int_{-\infty}^{v} \tilde{\varphi}_{X_1|Z_1}(Pv + H^T\sigma)\ \beta \operatorname{sgn}(H\Gamma\sigma) e^{-\beta|H\Gamma\sigma|}\Gamma^T e^{-\gamma\sqrt{\frac{\Gamma^T v}{H\Gamma} - \sigma + jz_2}\left(\frac{\Gamma^T v}{H\Gamma} - \sigma\right)}\, d\sigma, \tag{5.11}$$

where $P = I - H^T(H\Gamma)^{-1}\Gamma^T$. The second partial derivatives of $\bar{\varphi}_{X_2|Y_2}(v)$ with respect to v are obtained by differentiating (5.11). This requires taking the differential with respect to v of $e^{-\gamma\sqrt{\frac{\Gamma^T v}{H\Gamma} - \sigma + jz_2}\left(\frac{\Gamma^T v}{H\Gamma} - \sigma\right)}$, which is piecewise continuously differentiable and bounded, and of $\partial\tilde{\varphi}_{X_1|Z_1}(Pv + H^T\sigma)/\partial v$, which is continuously differentiable by the result of the previous subsection. This implies that the integrands in (5.11) are piecewise continuous functions of σ, are continuous in v, and have piecewise continuous and bounded first partial derivatives with respect to v. Consequently, this implies that the Hessian of $\bar{\varphi}_{X_2|Y_2}(v)$ is continuous [2].

5.1.3. Continuity of the First Two Derivatives of $\bar{\varphi}_{X_{k+1}|Y_{k+1}}$. In a recursive manner, the continuity of the first two derivatives of the characteristic function is now shown to be maintained when propagating from time step k to time step $k+1$, as it was maintained in going from times step 1 to 2 in subsection 5.1.2. It is assumed that using the recursion up to stage k, $\bar{\varphi}_{X_k|Y_k}(v)$ as given in (4.1) is twice continuously differentiable.

The time propagated characteristic function from k to $k+1$, given in (4.4), is written as

$$\bar{\varphi}_{X_{k+1}|Y_k}(v) = \bar{\varphi}_{X_k|Y_k}(\Phi^T v)\, e^{-\beta|\Gamma^T v|}. \tag{5.12}$$

Since $e^{-\beta|\Gamma^T v|}$ is not continuously differentiable, $\bar{\varphi}_{X_{k+1}|Y_k}(v)$ is not continuously differentiable and minimum variance estimates cannot be determined before the next measurement is processed.

The measurement updated characteristic function at $k+1$ is given in (4.10) and rewritten here using the explicit form of (5.12) as

$$\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{\varphi}_{X_k|Z_k}(v - H^T\eta) e^{-\beta|\Gamma^T v - H\Gamma\eta|} e^{-\gamma|\eta| + jz_{k+1}\eta}\, d\eta, \tag{5.13}$$

where $\tilde{\varphi}_{X_k|Y_k}(v) = \bar{\varphi}_{X_k|Y_k}(\Phi^T v)$. The first derivative of $\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v)$ is obtained in a manner similar to that used to construct (5.11). To take the second derivative, the change of variables given in (5.10) is used to obtain

$$\frac{\partial \bar{\varphi}_{X_{k+1}|Y_{k+1}}(v)}{\partial v} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\partial \tilde{\varphi}_{X_k|Y_K}(Pv + H^T\sigma)}{\partial v} e^{-\beta|H\Gamma\sigma|} e^{-\gamma\sqrt{\frac{\Gamma^T v}{H\Gamma} - \sigma + jz_{k+1}}\left(\frac{\Gamma^T v}{H\Gamma} - \sigma\right)}\, d\sigma$$
$$- \frac{1}{2\pi}\int_{-\infty}^{v} \tilde{\varphi}_{X_k|Y_k}(Pv + H^T\sigma)\ \beta \operatorname{sgn}(H\Gamma\sigma) e^{-\beta|H\Gamma\sigma|}\Gamma^T e^{-\gamma\sqrt{\frac{\Gamma^T v}{H\Gamma} - \sigma + jz_{k+1}}\left(\frac{\Gamma^T v}{H\Gamma} - \sigma\right)}\, d\sigma. \tag{5.14}$$

The integrands in (5.14) are at least piecewise continuous functions of σ and continuous in v. Moreover, their first partial derivatives with respect to v are piecewise continuous and bounded. Therefore, $\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v)$ is twice continuously differentiable [2], thereby justifying the supposition that the recursion produces a twice differentiable characteristic function after each measurement.

5.1.4. Measurement Matrix H with Zero Elements. Suppose H has only one zero element, $h_j = 0$, and assume that (H, Φ) is an observable pair. Therefore, (5.1) can be written as

$$\bar{\varphi}_{X_1|Z_1}(v) = \frac{e^{-\alpha_j|(e_j, v)|}}{2\pi}\int_{-\infty}^{\infty}\prod_{\substack{i=1\\ i\neq j}} e^{-\alpha_i|v_i - h_i\eta|} e^{-\gamma|\eta| + jz_1\eta}\, d\eta = \frac{e^{-\alpha_j|(e_j, v)|}}{2\pi}\hat{\varphi}_{X_1|Z_1}(\hat{e}_j, v), \tag{5.15}$$

where $\hat{\varphi}_{X_1|Z_1}(\hat{e}_j, v)$ is determined by the integral in (5.15) and $\hat{e}_j$ is a $n \times n-1$ matrix defined as $\hat{e}_j = \{e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n\}$. Due to our previous result in subsection 5.1.1, $\hat{\varphi}_{X_1|Z_1}(\hat{e}_j, v)$ and thus $\bar{\varphi}_{X_1|Z_1}(v)$

is twice continuously differentiable with respect to the $(\hat{e}_j, v) = \{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{n_x}\}^T$. However, clearly $\varphi_{X_1|Z_1}(v)$ is not continuously differentiable with respect to $(e_j, v) = v_j$ due to the term $e^{-\alpha_j|(e_j, v)|}$ in (5.15). Therefore, only a reduced-order state estimate can be constructed at $k = 1$.

Using (5.15), the time propagated characteristic function at $k = 2$ is given by

$$\bar{\varphi}_{X_2|Z_1}(v) = \sum_j \varphi_{X_1|Z_1}(\hat{e}_j^T \Phi^T v) e^{-\alpha_j|e_j^T \Phi^T v|} e^{-\beta|(\Gamma, v)|}. \tag{5.16}$$

After the second measurement update it is expressed as

$$\bar{\varphi}_{X_2|Y_2}(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\varphi}_{X_1|Z_1}(\Phi^T v - H\Phi \hat{e}_j \eta)\, e^{-\alpha_j|e_j^T \Phi^T v - H\Phi \hat{e}_j \eta)|} e^{-\beta|\Gamma^T (v - H^T \eta)|} e^{-\gamma|\eta| + jz_2\eta}\, d\eta. \tag{5.17}$$

If $H\Phi e_j = 0$, as in (5.15), (5.17) decomposes as

$$\bar{\varphi}_{X_2|Y_2}(v) = e^{-\alpha_j|e_j^T \Phi^T v|} \hat{\varphi}_{X_2|Y_2}(\hat{e}_j^T \Phi^T v), \tag{5.18}$$

where $\hat{\varphi}_{X_2|Y_2}(\hat{e}_j^T \Phi^T v)$ is defined by the remaining integral in (5.17) after removing $e^{-\alpha_j|e_j^T \Phi^T v|}$ from it. Clearly, in this case $\bar{\varphi}_{X_2|Y_2}(v)$ is not continuously differentiable at $v = \{0\}_n$. However, if $H\Phi e_j \neq 0$, there are no terms that can be factored out like in (5.18). Moreover, since the integrand in (5.17) is continuous in η and v, and has piecewise continuous and bounded derivatives with respect to those variables, using the procedure established in subsection 5.1.2, $\bar{\varphi}_{X_2|Y_2}(v)$ is twice continuously differentiable. Consequently, due to the result in subsection 5.1.3, when $H\Phi e_j \neq 0$, $\bar{\varphi}_{X_k|Y_k}(v)$ is twice continuously differentiable for any $k \geq 2$.

Now we return to the case that $H\Phi e_j = 0$. Using the steps presented above, if $H\Phi^2 e_j \neq 0$, then $\bar{\varphi}_{X_k|Y_k}(v)$ is twice continuously differentiable for any $k \geq 3$. Similarly, if also $H\Phi^2 e_j = 0$ but $H\Phi^3 e_j \neq 0$, continuity can be proved for $k \geq 4$. Since the system is assumed to be observable, there always exist $1 \leq \pounds \leq n$ such that $H\Phi^{p-1} e_j = 0 \; \forall p < \pounds$ and $H\Phi^{\pounds - 1} e_j \neq 0$. Consequently, $\bar{\varphi}_{X_k|Y_k}(v)$ is twice continuously differentiable for any $k \geq \pounds$.

This procedure can be extended to any number of zero elements of H.

5.2. Construction of the Conditional Mean and Estimation Error Variance. Having established in the previous section that $\bar{\varphi}_{X_k|Y_k}(v)$ is twice continuously differentiable, the explicit form of the pdf of the measurement history, the conditional mean, and conditional variance is determined from (4.1-4.3) by evaluating $\bar{\varphi}_{X_k|Y_k}(v)$ and its derivative as $v \to \{0\}_n$. We choose $v = \hat{v}$, where ε is a positive scalar such that $\varepsilon \to 0$ and $\hat{v}$ is a fixed direction in the v domain for which

$$(a_i^{k|k}, \hat{v}) \neq 0, \quad \forall \; (i, \pounds). \tag{5.19}$$

Along this direction the conditional mean and conditional variance will be shown to be easily computable, avoiding the discontinuity issues of $\operatorname{sgn}(a_i^{k|k}, \hat{v})$ when the condition in (5.19) does not hold. Due to the continuity of $\bar{\varphi}_{X_k|Y_k}(v)$ and its first two derivatives, any $\hat{v}$ can be chosen as long as (5.19) holds. With this choice of v,

$$\operatorname{sgn}(a_i^{k|k}, \hat{v}) = \operatorname{sgn}(a_i^{k|k}, \hat{v}) = s^i \tag{5.20}$$

are well defined constants that satisfy $|s^i| = 1$. From (4.1-4.3), $\bar{\varphi}_{X_k|Y_k}(\cdot)$ along this direction is given by

$$\bar{\varphi}_{X_k|Y_k}(\varepsilon \hat{v}) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(y_{gi}^{k|k}(\varepsilon \hat{v}), \exp(y_{ei}^{k|k}(\varepsilon \hat{v})), \tag{5.21}$$

where the functions $g_i^{k|k}(\cdot)$ are given in (4.2). Their arguments, defined in (4.3), are

$$y_{gi}^{k|k}(\varepsilon \hat{v}) = \sum_{=1}^{n_{ei}^{k|k}} q_i^{k|k} \operatorname{sgn}(a_i^{k|k}, \hat{v}) = \sum_{=1}^{n_{ei}^{k|k}} q_i^{k|k} s^i = y_{gi}^{k|k}(\hat{v}) \in \mathbb{R}^k. \tag{5.22}$$

Since the vector $\hat{v}$ is constant, the arguments $y_{gi}^{k|k}(\cdot)$ and thus the coefficient functions $g_i^{k|k}(\cdot)$ are constant along the chosen direction. This will simplify greatly the evaluation of $\bar{\varphi}_{X_k|Y_k}(\cdot)$ and its derivatives at the origin of the spectral variable v.

Similarly, the arguments $y_{ei}^{k|k}(\cdot)$ of the exponents in (5.21), defined in (4.3), are manipulated as follows

$$y_{ei}^{k|k}(\hat{v}) = -\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k} \left| (a_i^{k|k}, \hat{v}) \right| + j(b_i^{k|k}, \hat{v}) = -\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k} \operatorname{sgn}(a_i^{k|k}, \hat{v}) (a_i^{k|k}, \hat{v}) + j(b_i^{k|k}, \hat{v}) \tag{5.23}$$

$$= \left[-\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k} s^i (a_i^{k|k}, \hat{v}) + j(b_i^{k|k}, \hat{v}) \right] = (\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v}),$$

where we have defined

$$\bar{y}_{ei}^{k|k}(\hat{v}) = -\sum_{i=1}^{n_{ei}^{k|k}} p_i^{k|k} s^i a_i^{k|k} + j b_i^{k|k}, \tag{5.24}$$

which is a n-dimensional vector. Since the vector $\hat{v}$ is constant, so is the expression $(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})$, making $y_{ei}^{k|k}(\cdot)$ of the exponents linear in ρ along the chosen direction. Using the above results, (5.21) can be restated as

$$\bar{\varphi}_{X_k|Y_k}(\rho \hat{v}) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(\rho, y_{gi}^{k|k}(\hat{v})) \exp(\rho (\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})). \tag{5.25}$$

This last form of $\bar{\varphi}_{X_k|Y_k}(\rho \hat{v})$ will now be used to compute the minimum variance estimate and its error variance.

The conditional mean of the state x_k given the data sequence y_k is given by

$$\hat{x}_k = E[x_k | y_k] = \frac{1}{j f_{Y_k}(y_k)} \left(\frac{\partial \bar{\varphi}_{X_k|Y_k}(\rho \hat{v})}{\partial(\rho \hat{v})} \right)^T \Bigg|_{\rho=0}. \tag{5.26}$$

The pdf $f_{Y_k}(y_k)$ needed to normalize the above result is determined by evaluating $\bar{\varphi}_{X_k|Y_k}(\rho \hat{v})$ at $\rho = 0$. Using (5.25) we obtain that

$$f_{Y_k}(y_k) = \bar{\varphi}_{X_k|Y_k}(\rho \hat{v})\Big|_{\rho=0} = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(0, y_{gi}^{k|k}(\hat{v})), \tag{5.27}$$

where $g_i^{k|k}(\cdot)$ are given in (4.2) and the arguments $y_{gi}^{k|k}(\hat{v})$ were defined in (5.22).

The derivative of $\bar{\varphi}_{X_k|Y_k}(\cdot)$ used in (5.26) is determined by differentiating (5.25). Since, as stated earlier, $g_i^{k|k}(\rho, y_{gi}^{k|k}(\hat{v}))$ and $(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})$ are constant, the derivative is evaluated as

$$\left(\frac{\partial \bar{\varphi}_{X_k|Y_k}(\rho \hat{v})}{\partial(\rho \hat{v})} \right)^T = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(\rho, y_{gi}^{k|k}(\hat{v})) \bar{y}_{ei}^{k|k}(\hat{v}) \exp(\rho (\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})). \tag{5.28}$$

Evaluating (5.28) at $\rho = 0$ yields the minimum conditional variance estimate

$$\hat{x}_k = \frac{1}{j f_{Y_k}(y_k)} \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(0, y_{gi}^{k|k}(\hat{v})) \bar{y}_{ei}^{k|k}(\hat{v}), \tag{5.29}$$

where $f_{Y_k}(y_k)$, $g_i^{k|k}(\cdot)$, $y_{gi}^{k|k}(\hat{v})$, and $\bar{y}_{ei}^{k|k}(\hat{v})$ are given in (5.27), (4.2), (5.22) and (5.24), respectively.

The second conditional moment of x_k given y_k is determined by

$$E[x_k x_k^T | y_k] = \frac{1}{j^2 f_{Y_k}(y_k)} \left. \frac{\partial^2 \bar{\varphi}_{X_k|Y_k}(\rho \hat{v})}{\partial(\rho \hat{v}) \partial(\rho \hat{v})^T} \right|_{\rho=0}. \tag{5.30}$$

The second derivative in the above equation is computed by differentiating (5.28) to yield

$$\frac{\partial^2 \bar{\varphi}_{X_k|Y_k}(\rho \hat{v})}{\partial(\rho \hat{v}) \partial(\rho \hat{v})^T} = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(\rho, y_{gi}^{k|k}(\hat{v})) \bar{y}_{ei}^{k|k}(\hat{v}) \bar{y}_{ei}^{k|k}(\hat{v})^T \exp(\rho (\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})). \tag{5.31}$$

Hence the second moment is given by

$$E[x_k x_k^T | y_k] = \frac{1}{j^2 f_{Y_k}(y_k)} \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}(\hat{v}) \begin{bmatrix} y_{gi}^{k|k}(\hat{v}) \\ \bar{y}_{ei}^{k|k}(\hat{v}) \end{bmatrix} \begin{bmatrix} y_{gi}^{k|k}(\hat{v}) \\ \bar{y}_{ei}^{k|k}(\hat{v}) \end{bmatrix}^T. \tag{5.32}$$

Finally, from (5.29) and (5.32), the error $\tilde{x}_k = x_k - \hat{x}_k$ variance can be evaluated as

$$E[\tilde{x}_k \tilde{x}_k^T | y_k] = E[x_k x_k^T | y_k] - \hat{x}_k \hat{x}_k^T. \tag{5.33}$$

5.3. Conditional Mean and Estimation Error Variance at $k = 1$. To appreciate the complex results given in section 5.2 and to understand the underlying structure of the Cauchy estimator, greatly simplified and compact expressions are determined for $k = 1$. Here we address the case that $h_i \neq 0$ for $i = \{1, \ldots, n\}$. The characteristic function $\bar{\varphi}_{X_k|Z_k}(v)$ evaluated along $v = \hat{v}$ is given in (5.25). For $k = 1$ it can be easily simplified to

$$\bar{\varphi}_{X_1|Z_1}(\hat{v}) = \frac{1}{2\pi} \prod_{i=1}^{n+1} \left[jz_1 + \rho_i + \sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l s_l \right]^{-1} - \prod_{i=1}^{n+1} \left[-jz_1 - \rho_i + \sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l s_l \right]^{-1} \exp\left(-\sum_{i=1}^{n+1} \rho_i (a_i^T \hat{v}) s_i + jz_1 \mu_{n+1}^T \hat{v} \right). \tag{5.34}$$

The pdf of the measurement history, the conditional mean, and the conditional variance are found by evaluating (5.34) and its first two derivatives at $= 0$. There are many straight forward manipulations needed to obtain our final compact results presented here. These detailed derivations are omitted here for brevity.

First, $f_{Y_k}(y_k)_{k=1} = f_{Z_1}(z_1)$, given in (5.27), is determined by letting $\to 0$ in (5.34) and noting that most of the term in the remaining sum cancel leaving all but the first and last terms. The final form is thus determined as

$$f_{Z_1}(z_1) = \frac{1}{2\pi} \left(\prod_{i=2}^{n+1} \left[jz_1 + \rho_1 + \rho_i \right]^{-1} - \prod_{i=1}^{n+1} \left[-jz_1 - \rho_i \right]^{-1} \right) = \frac{1}{\pi} \frac{\sum_{i=1}^{n} \alpha_i |h_i| + \gamma}{z_1^2 + \left(\sum_{i=1}^{n} \alpha_i |h_i| + \gamma \right)^2}. \tag{5.35}$$

In the last equality, the initial parameters in (2.4) and (3.11) are used.

The conditional mean is computed by taking the derivative of $\bar{\varphi}_{X_1|Z_1}(\hat{v})$ given by (5.34) and then letting $\to 0$. We obtain

$$\frac{\partial \bar{\varphi}_{X_1|Z_1}(\hat{v})}{\partial \hat{v}}\bigg|_{E=0} = \frac{1}{2\pi} \sum_{i=1}^{n+1} \left[jz_1 + \rho_i + \sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l s_l \right]^{-1} - \left[jz_1 - \rho_i + \sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l s_l \right]^{-1} \left[\sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l s_l a_i^T + jz_1 \mu_i^T \right], \tag{5.36}$$

where $a_i = -a_i$ was used in (5.36). To simplify the above, place numerator terms over the common denominator elements as given in the sum of (5.36). In doing so, after a bit of straight forward algebra and using the decomposition $a_i = \mu_i - \mu$, these elements simplify such that all but the first and last terms cancel. Therefore, we obtain

$$\frac{\partial \bar{\varphi}_{X_1|Z_1}(\hat{v})}{\partial \hat{v}}\bigg|_{E=0} = \frac{1}{\pi} \frac{\left(jz_1 - \sum_{i=1}^{n+1} \rho_i \mu_i^T \right)}{z_1^2 + \left(\sum_{i=1}^{n+1} \rho_i \right)^2}. \tag{5.37}$$

Dividing (5.37) by (5.35) and the imaginary number j, the conditional mean is determined as

$$\hat{x}_1 = \frac{z_1 - \sum_{i=1}^{n+1} \rho_i \mu_i}{\sum_{i=1}^{n+1} \rho_i} = z_1 \frac{\left[\alpha_1 \operatorname{sgn}(h_1) \quad \cdots \quad \alpha_n \operatorname{sgn}(h_n) \right]^T}{\sum_{i=1}^{n} \alpha_i |h_i| + \gamma}. \tag{5.38}$$

This is the explicit form of the conditional mean given in (5.29) evaluated and simplified at $k = 1$.

The conditional second moment is determined from the second partial derivative of $\bar{\varphi}_{X_1|Z_1}(\hat{v})$ with respect to $\hat{v}$ and letting $\varepsilon \to 0$ as

$$\frac{\partial^2 \bar{\varphi}_{X_1|Z_1}(\hat{v})}{\partial \hat{v} \partial \hat{v}^T}\bigg|_{\varepsilon=0} = \frac{1}{2\pi}\sum_{i=1}^{n+1}\left[jz_1 + \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n+1}\rho_j s_j^{-1}\left(-jz_1 - \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n+1}\rho_j s_j^{-1}\right)\right] \times \left(-\sum_{\substack{j=1 \\ j\neq i}}^{n+1}\rho_j s_j a_j + jz_1\mu_i\right)\left(-\sum_{\substack{j=1 \\ j\neq i}}^{n+1}\rho_j s_j a_j^T + jz_1\mu_i^T\right) \tag{5.39}$$

To simplify (5.39), place numerator terms over the common denominator elements as given explicitly in the sum in (5.39). In doing so, after a bit of straight forward algebra, having made use of the decomposition $a_i = \mu_i - \mu$, these elements simplify greatly. The result is

$$\pi\frac{\partial^2 \bar{\varphi}_{X_1|Z_1}(\hat{v})}{\partial \hat{v} \partial \hat{v}^T}\bigg|_{\varepsilon=0} = \frac{\left(\sum_{j=1}^{n+1}\rho_j\right)\left(\sum_{j=1}^{n+1}\rho_j\mu_j\right)\left(\sum_{j=1}^{n+1}\rho_j\mu_j^T\right)}{z_1^2\left(\sum_{j=1}^{n+1}\rho_j\right)^2} - \sum_{j=1}^{n}\rho_j\mu_j\mu_j^T. \tag{5.40}$$

The conditional second moment is given by

$$E\left[x_1 x_1^T | z_1\right] = -\frac{1}{f_{Z_1}(z^1)}\cdot\frac{\partial^2 \bar{\varphi}_{X_1|Z_1}(\hat{v})}{\partial \hat{v} \partial \hat{v}^T}\bigg|_{\varepsilon=0}, \tag{5.41}$$

where $f_{Z_1}(z_1)$ is found in (5.35). The conditional error variance, where $\tilde{x}_1 = x_1 - \hat{x}_1$, is then determined as

$$E\left[\tilde{x}_1\tilde{x}_1^T | z_1\right] = \left(1 + \frac{z_1^2}{\left(\sum_{i=1}^{n}\alpha_i|h_i| + \gamma\right)^2}\right)\begin{bmatrix} \left(\sum_{i=2}^{n}\alpha_i|h_i| + \gamma\right) & \cdots & -\alpha_1\alpha_n \operatorname{sgn}(h_1)\operatorname{sgn}(h_n) \\ \vdots & & \vdots \\ -\alpha_1\alpha_n \operatorname{sgn}(h_1)\operatorname{sgn}(h_n) & \cdots & \left(\sum_{i=1}^{n-1}\alpha_i|h_i| + \gamma\right)\frac{\alpha_n}{|h_n|} \end{bmatrix}. \tag{5.42}$$

This is an explicit form of the conditional variance given in (5.33) for $k = 1$.

It is surprising that at $k = 1$ the pdf of the measurement history, the minimum variance estimate and the estimation error covariance matrix given, respectively, in (5.35), (5.38), and (5.42) can be expressed analytically in such a compact and relatively simple form as functions of the system parameters and the first measurement. Such simplicity is not retained for the subsequent time steps. It is also interesting to note that, contrary to the Gaussian case where the estimation error variance is known *a priori*, the estimation error variance of (5.42) is an explicit (quadratic) function of the measurement z_1.

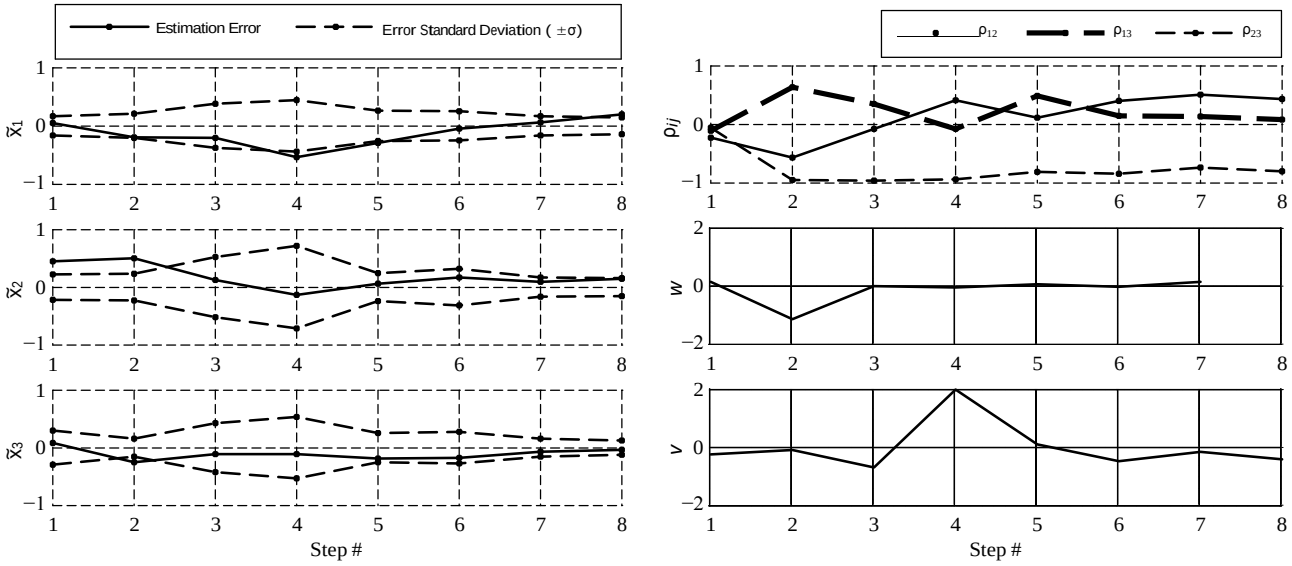
6. Three-Dimensional State Estimator. The performance of the proposed estimator was tested numerically. The system parameters were chosen as follows:

$$\Phi = \begin{bmatrix} 1.4 & -0.6 & -1.0 \\ -0.2 & 1.0 & 0.5 \\ 0.6 & -0.6 & -0.2 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0.1 \\ 0.3 \\ -0.2 \end{bmatrix}, \quad \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.08 \\ 0.05 \end{bmatrix},$$

$$H = \begin{bmatrix} 1.0 & 0.5 & 0.2 \end{bmatrix}, \quad \beta = 0.1, \quad \gamma = 0.2.$$

The system has stable eigenvalues at $0.7 \pm 0.3j$ and 0.8. It is observable and complies with the necessary condition that $H\Gamma \neq 0$.

The simulation results are depicted in Fig. 6.1, showing the estimation errors in 6.1(a), and the cross-correlation, process and measurement noises in 6.1(b). It can be clearly observed that the estimation error standard deviation depends on the measurement sequence. Specifically, the standard deviation increases when a large measurement noise is encountered, e.g., at time step #4.



(a) Estimation Errors (b) Cross-correlations and system noises

Fig. 6.1. *Three-State Estimation Example: Numerical Results*

7. Conclusions. For a linear discrete-time system with additive Cauchy measurement and process noises, an analytic recursion of the characteristic function of the unnormalized condition pdf is determined for the n-vector state estimator. This characteristic function is shown to be a growing sum of terms involving a coefficient function multiplied by an exponential whose arguments are weighted nonlinear functions of the spectral variable. It is then shown that this characteristic function is twice continuously differentiable, allowing the determination of the conditional mean and the conditional second moment, from which the conditional error variance is determined. These moments are a nonlinear function of the measurement history, where the conditional mean constitutes the Cauchy estimator for the multivariable system. The estimator was then evaluated numerically for a third-order example. At this stage of the numerical development, only a few steps in time are shown for this example because of the excessively large number of terms needed to represent the characteristic function. Nonetheless, this is the first time that a closed-form analytical estimator is derived for a multivariable system with heavy tailed Cauchy process and measurement noises.

REFERENCES

[1] G. R. Arce, *Nonlinear Signal Processing: A statistical Approach*, Wiley, New Jersey, 2005.
[2] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Wiley-Interscience, 1984.
[3] M. Idan and J. L. Speyer, *Cauchy estimation for linear scalar systems*, IEEE Transactions on Automatic Control, 55 (2010), pp. 1329–1342.
[4] ———, *State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach*, SIAM J. Control Optim., 0 (2012), pp. 1–24.
[5] L. N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons, Inc., second ed., 1994, ch. 16.
[6] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, *Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution*, IEEE Communications Letters, 2 (1998), pp. 282–284.
[7] C. J. Masreliez and R. D. Martin, *Robust bayesian estimation for the linear model and robustifying the kalman filter*, IEEE Transactions on Automatic Control, AC-22 (1977), pp. 361–371.
[8] J. P. Nolan, J. G. Gonzalez, and Ná nez, *Stable filters: A robust signal processing framework for heavy-tailed noise.*, in In Proceedings of the 2010 IEEE Radar Conference, 2010.
[9] P. Reeves, *A non-gaussian turbulence simulation*, Tech. Rep. AFFDL-TR-69-67, Air Force Flight Dynamics Laboratory, 1969.
[10] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall, New York, 1994.
[11] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control*, SIAM, 2008.
[12] Nassim N. Taleb, *The Black Swan: The Impact of the Highly Improbable*, Random House, 2007.
[13] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhauser, 1998.

Appendix

Appendix A. Characteristic Function of an Unnormalized Conditional PDF Given a Scalar Measurement. Consider a scalar measurement of an n-dimensional state vector as

$$z = Hx + v = \sum_{i=1}^{n} h_i x_i + v. \tag{A.1}$$

where z is a given scalar measurement, $x \in \mathbb{R}^n$, and v is Cauchy distributed. For this measurement, the characteristic function of the unnormalized conditional density function is given as

$$\bar{\varphi}_{x|z}(\nu) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x) f_V(z - Hx) e^{j\nu^T x} dx. \tag{A.2}$$

The above resembles a Fourier transform of a product of two functions: $f_X(x)$ and $f_V(z - Hx)$. Using the dual convolution property, this integral can be solved by a convolution in the ν domain between the characteristic function of $f_X(x)$, i.e., $\varphi_X(\nu)$, and the characteristic function of $f_V(z - Hx)$, which we denote by $\hat{\varphi}_V(\nu)$. Hence, $\bar{\varphi}_{x|z}(\nu)$ can be computed by the convolution integral

$$\bar{\varphi}_{x|z}(\nu) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \varphi_X(\nu - \eta) \hat{\varphi}_V(\eta) d\eta. \tag{A.3}$$

The characteristic function $\hat{\varphi}_V(\nu)$ is determined as follows

$$\hat{\varphi}_V(\nu) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V(z - Hx) e^{j\nu^T x} dx = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V\left(z - \sum_{i=1}^{n} h_i x_i\right) e^{j\left(\sum_{i=1}^{n} \nu_i x_i\right)} dx_1 \cdots dx_n. \tag{A.4}$$

Assuming, without loss of generality, that $h_n \neq 0$, we perform the change of variables $\xi = z - \sum_{i=1}^{n} h_i x_i$. This implies $x_n = \left(z - \xi - \sum_{i=1}^{n-1} h_i x_i\right)/h_n$ and for fixed $x_i, i = 1, \ldots, n-1$, $dx_n = d\xi/|h_n|$. Using this change of variables, (A.4) is manipulated as

$$\begin{aligned}
\hat{\varphi}_V(\nu) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V(\xi) e^{j\left(\sum_{i=1}^{n-1} \nu_i x_i + \nu_n \frac{z - \xi - \sum_{i=1}^{n-1} h_i x_i}{h_n}\right)} d\xi/|h_n| dx_1 \cdots dx_{n-1} \\
&= \frac{e^{j\nu_n z/h_n}}{|h_n|} \int_{-\infty}^{\infty} f_V(\xi) e^{-j\frac{\nu_n}{h_n}\xi} d\xi \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{j\sum_{i=1}^{n-1}\left(\nu_i - \frac{h_i}{h_n}\nu_n\right)x_i} dx_1 \cdots dx_{n-1} \\
&= \frac{e^{j\nu_n z/h_n}}{|h_n|} \varphi_V\left(-\frac{\nu_n}{h_n}\right) \prod_{i=1}^{n-1} \int_{-\infty}^{\infty} e^{j\left(\nu_i - \frac{h_i}{h_n}\nu_n\right)x_i} dx_i = \frac{e^{j\nu_n z/h_n}}{|h_n|} \varphi_V\left(-\frac{\nu_n}{h_n}\right) \prod_{i=1}^{n-1} (2\pi)\delta\left(\nu_i - \frac{h_i}{h_n}\nu_n\right),
\end{aligned} \tag{A.5}$$

where $\delta(\cdot)$ is the Dirac delta function. Substitute (A.5) into (A.3) to obtain

$$\begin{aligned}
\bar{\varphi}_{x|z}(\nu) &= \frac{(2\pi)^{n-1}}{(2\pi)^n |h_n|} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \varphi_X(\nu - \eta) e^{j\frac{\eta_n}{h_n}z} \varphi_V\left(-\frac{\eta_n}{h_n}\right) \prod_{i=1}^{n-1} \delta\left(\eta_i - \frac{h_i}{h_n}\eta_n\right) d\eta_1 \cdots d\eta_{n-1} d\eta_n \\
&= \frac{1}{2\pi|h_n|} \int_{-\infty}^{\infty} \varphi_X\left(\nu - H^T \frac{\eta_n}{h_n}\right) e^{j\frac{\eta_n}{h_n}z} \varphi_V\left(-\frac{\eta_n}{h_n}\right) d\eta_n.
\end{aligned} \tag{A.6}$$

Finally, using the change of variables $\eta = \eta_n/h_n$ and hence $d\eta_n = d\eta|h_n|$, the last equation can be restated as

$$\bar{\varphi}_{x|z}(\nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(\nu - H^T\eta) \varphi_V(-\eta) e^{jz\eta} d\eta. \tag{A.7}$$

Appendix B. Integral of an Exponent of Absolute Values. The measurement update stage of the Cauchy estimator entails evaluating a convolution integral that involves an exponential of absolute values of affine functions in the spectral variables. In this appendix we present a methodology for carrying out such integrals.

B.1. Exponent-Only Integral. In the first scalar measurement update, e.g., (3.8), we encountered the basic integral of the form

$$I = \int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n} \rho_i |\xi_i - \eta| + jz\eta\right) d\eta, \tag{B.1}$$

where z is the measurement, ρ_i-s are positive constants, and the ξ_i-s are variables linear in v. The difficulty in evaluating this integral is that the ξ_i-s must be ordered to carry out the integration. However, since the ξ_i-s are functions of the spectral variables v-s which take on any value, any ordering of ξ_i-s is possible. We show that the solution to this integral can be expressed in closed form regardless of the order. For convenience, define

$$\xi_0 = -\infty, \qquad \xi_{n+1} = \infty. \tag{B.2}$$

To carry out the integral, we first assume that $\xi_\ell \le \xi_i$ for all $\ell < i$ with $(\ell, i) \in \{0, \ldots, n+1\}$. In this case the integral (B.1) can be decomposed into the sum

$$I = \sum_{i=0}^{n} \int_{\xi_i}^{\xi_{i+1}} \exp\left(-\sum_{j=1}^{n} \rho_j |\xi_j - \eta| + jz\eta\right) d\eta = \sum_{i=0}^{n} \int_{\xi_i}^{\xi_{i+1}} \exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \eta)\operatorname{sgn}(\xi_j - \eta) + jz\eta\right) d\eta. \tag{B.3}$$

Note that $\operatorname{sgn}(\xi_j - \eta)$ is constant over the interval $\eta \in (\xi_i, \xi_{i+1}), \forall i$. This constant depends on the relative size of the indices ℓ and i. With this in mind, for η in the closure $\eta \in [\xi_i, \xi_{i+1}]$, we can define the sign function in (B.3) as

$$\operatorname{sgn}(\xi_j - \eta) = s_{ji} = \begin{cases} \operatorname{sgn}(\xi_j - \xi_i) & \text{if } i \neq \ell \\ -1 & \text{if } i = \ell \end{cases} \tag{B.4}$$

From (B.4), for a given i, the discrete, two indexed function s_{ji} is constant for all ℓ except for one switch at $\ell = i$, i.e., $s_{ji} = -1\ \forall\ \ell \le i$ and $s_{ji} = 1\ \forall\ \ell > i$. Moreover, it can be concluded that $s_{ji} = s_{j,i-1}\ \forall\ i \neq \ell$.

Using the definitions in (B.4), (B.3) can be restated in terms of s_{ji} and integrated as

$$\begin{aligned} I &= \sum_{i=0}^{n} \int_{\xi_i}^{\xi_{i+1}} \exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \eta) s_{ji} + jz\eta\right) d\eta \\ &= \sum_{i=0}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \eta) s_{ji} + jz\eta\right)\Big|_{\xi_i}^{\xi_{i+1}}}{jz + \sum_{j=1}^{n} \rho_j s_{ji}} \\ &= \sum_{i=0}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_{i+1}) s_{ji} + jz\xi_{i+1}\right)}{jz + \sum_{j=1}^{n} \rho_j s_{ji}} - \sum_{i=0}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_i) s_{ji} + jz\xi_i\right)}{jz + \sum_{j=1}^{n} \rho_j s_{ji}}. \end{aligned} \tag{B.5}$$

The first sum in (B.5) can be manipulated as follows

$$\begin{aligned} &\sum_{i=0}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_{i+1}) s_{ji} + jz\xi_{i+1}\right)}{jz + \sum_{j=1}^{n} \rho_j s_{ji}} \\ &= \sum_{i=0}^{n-1} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_{i+1}) s_{ji} + jz\xi_{i+1}\right)}{jz + \sum_{j=1}^{n} \rho_j s_{ji}} + \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_{n+1}) s_{jn} + jz\xi_{n+1}\right)}{jz + \sum_{j=1}^{n} \rho_j s_{jn}}, \\ &\qquad \underbrace{\qquad\qquad}_{=0 \text{ because } s_{ji}=-1\ \forall\ j \le n \text{ and } \xi_{n+1}=\infty} \\ &= \sum_{i=1}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_i) s_{j,i-1} + jz\xi_i\right)}{jz + \sum_{j=1}^{n} \rho_j s_{j,i-1}}. \end{aligned} \tag{B.6}$$

Since $s^i_{i-1} = 1$ and $s_{i-1} = s_i\ \forall\ i \neq \ell$, the above result can be simplified as

$$\sum_{i=1}^{n} \frac{\exp\left(-\sum_{j=1}^{n} \rho_j (\xi_j - \xi_i) s_{j,i-1} + jz\xi_i\right)}{jz + \sum_{j=1}^{n} \rho_j s_{j,i-1}} = \sum_{i=1}^{n} \frac{\exp\left(-\sum_{\substack{j=1 \\ j \neq i}}^{n} \rho_j (\xi_j - \xi_i) s_{ji} + jz\xi_i\right)}{jz + \rho_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \rho_j s_{ji}}. \tag{B.7}$$

In the above, the zero term $\rho_i(\xi_i - \xi_i)s^i_{i-1} = 0$ was dropped from the sum in the exponent.

The second sum in (B.5) can be manipulated in a similar fashion to yield

$$\text{(B.8)}\qquad \frac{\exp\left(-\sum_{i=1}^{n}\rho_i(\xi_i - \xi_i)s_i + jz\xi_i\right)}{\prod_{i=0}^{n}\left(jz + \sum_{i=1}^{n}\rho_i s_i\right)} = \prod_{i=1}^{n}\frac{\exp\left(-\sum_{i=1,\neq i}^{n}\rho_i(\xi_i - \xi_i)s_i + jz\xi_i\right)}{\left(jz - \rho_i + \sum_{i=1,\neq i}^{n}\rho_i s_i\right)}.$$

Substituting (B.7), (B.8), and $s_i = \operatorname{sgn}(\xi_i - \xi_i)\ \forall\ i \neq i$ of (B.4) into (B.5), the solution to (B.1) is expressed as

$$\text{(B.9)}\qquad I = \sum_{i=1}^{n} g_i\left(\sum_{i=1,\neq i}^{n}\rho_i \operatorname{sgn}(\xi_i - \xi_i)\right)\exp\left(-\sum_{i=1,\neq i}^{n}\rho_i|\xi_i - \xi_i| + jz\xi_i\right),$$

where

$$\text{(B.10)}\qquad g_i\left(\sum_{i=1,\neq i}^{n}\rho_i \operatorname{sgn}(\xi_i - \xi_i)\right) = \frac{1}{jz + \rho_i + \sum_{i=1,\neq i}^{n}\rho_i \operatorname{sgn}(\xi_i - \xi_i)} - \frac{1}{jz - \rho_i + \sum_{i=1,\neq i}^{n}\rho_i \operatorname{sgn}(\xi_i - \xi_i)}.$$

Changing the ordering of the ξ_i-s caused by changes in the v_i-s does not change the solution of the integral given by (B.10). Effectively, the sgn function in the exponential and denominator make the integral independent of the ordering of the ξ_i-s and thus giving a general result for the integral of (B.1).

B.2. Generalized Integral. In the second and subsequent measurement update steps we encounter a more general integral of the form

$$\text{(B.11)}\qquad I = \int_{-\infty}^{\infty} g\left(\sum_{i=1}^{n} Q_i \operatorname{sgn}(\xi_i - \eta)\right)\exp\left(-\sum_{i=1}^{n}\rho_i|\xi_i - \eta| + jz\eta\right)d\eta,$$

where Q_i could be m-dimensional vectors. We use the same methodology as before to solve the integral in (B.11). Using the definitions of ξ_0 and ξ_{n+1} in (B.2), and those of s_i in (B.4), the integral is restated as

$$\text{(B.12)}\qquad \begin{aligned} I &= \sum_{i=0}^{n}\int_{\xi_i}^{\xi_{i+1}} g\left(\sum_{i=1}^{n} Q_i \operatorname{sgn}(\xi_i - \eta)\right)\exp\left(-\sum_{i=1}^{n}\rho_i(\xi_i - \eta)\operatorname{sgn}(\xi_i - \eta) + jz\eta\right)d\eta \\ &= \sum_{i=0}^{n}\int_{\xi_i}^{\xi_{i+1}} g\left(\sum_{i=1}^{n} Q_i s_i\right)\exp\left(-\sum_{i=1}^{n}\rho_i(\xi_i - \eta)s_i + jz\eta\right)d\eta. \end{aligned}$$

It is integrated as

$$\text{(B.13)}\qquad \begin{aligned} I = &\sum_{i=0}^{n}\frac{g\left(\sum_{i=1}^{n} Q_i s_i\right)\exp\left(-\sum_{i=1}^{n}\rho_i(\xi_i - \xi_{i+1})s_i + jz\xi_{i+1}\right)}{\left(jz + \sum_{i=1}^{n}\rho_i s_i\right)} \\ &- \sum_{i=0}^{n}\frac{g\left(\sum_{i=1}^{n} Q_i s_i\right)\exp\left(-\sum_{i=1}^{n}\rho_i(\xi_i - \xi_i)s_i + jz\xi_i\right)}{\left(jz + \sum_{i=1}^{n}\rho_i s_i\right)} \end{aligned}$$

and manipulated as before to yield

$$I = \sum_{i=1}^{n} \left[g\left(Q_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j s_j\right) - g\left(-Q_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j s_j\right) \right] \frac{\exp\left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j (\xi_j - \xi_i) s_i\right) + jz\xi_i}{jz + \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j s_j - \left(jz - \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j s_j\right)}. \tag{B.14}$$

Substituting $s_j = \operatorname{sgn}(\xi_j - \xi_i) \;\forall\, j \neq i$, the integral becomes

$$I = \sum_{i=1}^{n} G_i\left(\sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j \operatorname{sgn}(\xi_j - \xi_i), \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j \operatorname{sgn}(\xi_j - \xi_i)\right) \exp\left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j |\xi_j - \xi_i| + jz\xi_i\right), \tag{B.15}$$

where

$$G_i\left(\sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j \operatorname{sgn}(\xi_j - \xi_i), \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j \operatorname{sgn}(\xi_j - \xi_i)\right) = \frac{g\left(Q_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j \operatorname{sgn}(\xi_j - \xi_i)\right) - g\left(-Q_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} Q_j \operatorname{sgn}(\xi_j - \xi_i)\right)}{jz + \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j \operatorname{sgn}(\xi_j - \xi_i) - \left(jz - \rho_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} \rho_j \operatorname{sgn}(\xi_j - \xi_i)\right)}. \tag{B.16}$$

Appendix C. Time Propagation of Characteristic Functions for Linear Systems. In this appendix we consider a multi-stage propagation of a random vector through a linear system, i.e. $x_{k+1} = \Phi x_k + \Gamma w_k$, where $x_k \in \mathbb{R}^n$ is the system state, the initial conditions $x_1 \in \mathbb{R}^n$ and process noise $w_k \in \mathbb{R}^m$ are independent random vectors with given pdfs $f_{X_1}(x_1)$ and $f_W(w)$, and characteristic functions $\varphi_{X_1}(v_x)$ and $\varphi_W(v_w)$, respectively. Note that here $v_x \in \mathbb{R}^n$, while $v_w \in \mathbb{R}^m$. For simplicity, the pdf and hence the characteristic function of w are assumed to be constant for all k. $\Phi \in \mathbb{R}^{n\times n}$ and $\Gamma \in \mathbb{R}^{n\times m}$ are known matrices, with $|\Phi| \neq 0$.

Assume that at step k the state pdf and characteristic function are given by $f_{X_k}(x_k)$ and $\varphi_{X_k}(v_x)$. The pdf of x_{k+1} is determined by using the linear transformation

$$\begin{pmatrix} x_{k+1} \\ w_k \end{pmatrix} = \underbrace{\begin{bmatrix} \Phi & \Gamma \\ 0 & I \end{bmatrix}}_{J^{-1}} \begin{pmatrix} x_k \\ w_k \end{pmatrix} \quad \leftrightarrow \quad \begin{pmatrix} x_k \\ w_k \end{pmatrix} = \underbrace{\begin{bmatrix} \Phi^{-1} & -\Phi^{-1}\Gamma \\ 0 & I \end{bmatrix}}_{|J|=|\Phi^{-1}|} \begin{pmatrix} x_{k+1} \\ w_k \end{pmatrix} \tag{C.1}$$

in the joint density $f_{X_k W}(x_k, w_k) = f_{X_k}(x_k) f_W(w_k)$ and finding the marginal density [11] as

$$f_{X_{k+1}}(x_{k+1}) = |\Phi^{-1}| \int_{-\infty}^{\infty} f_{X_k}(\Phi^{-1}x_{k+1} - \Phi^{-1}\Gamma w_k) f_W(w_k)\, dw_k. \tag{C.2}$$

The characteristic function of x_{k+1} is given by

$$\varphi_{X_{k+1}}(v) = |\Phi^{-1}| \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_{X_k}(\Phi^{-1}x_{k+1} - \Phi^{-1}\Gamma w_k) f_W(w_k)\, dw_k \right] e^{jv^T x_{k+1}} dx_{k+1} \tag{C.3}$$

Interchanging the order of integration, and using the substitution $x_{k+1} = \Phi x_k + \Gamma w_k \;\Rightarrow\; dx_{k+1} = |\Phi| dx_k$, the integral in (C.3) is solved as

$$\begin{aligned} \varphi_{X_{k+1}}(v) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_{X_k}(x_k)\, e^{jv^T(\Phi x_k + \Gamma w_k)} dx_k \right] f_W(w_k)\, dw_k \\ &= \left[\int_{-\infty}^{\infty} f_{X_k}(x_k)\, e^{jv^T \Phi x_k} dx_k \right] \left[\int_{-\infty}^{\infty} f_W(w_k)\, e^{jv^T \Gamma w_k} dw_k \right] = \varphi_{X_k}(\Phi^T v)\, \varphi_W(\Gamma^T v). \end{aligned} \tag{C.4}$$

B Appendix

Moshe Idan and Jason L. Speyer
Multivariate Cauchy Estimator with Scalar Measurement and Process Noises
Precedings of the 2013 IEEE Conference on Decision and Control

Multivariate Cauchy Estimator with Scalar Measurement and Process Noises

Moshe Idan
Aerospace Engineering
Technion - Israel Institute of Technology
Haifa, Israel, 32000
Email: moshe.idan@technion.ac.il

Jason L. Speyer
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: speyer@seas.ucla.edu

***Abstract*—The conditional mean estimator for a n-state linear system with additive Cauchy measurement and process noises is developed. For the multi-variable system state, the characteristic function of the unnormalized conditional probability density function is sequentially propagated through measurement updates and dynamic state propagation, while expressing the resulting characteristic function in a closed analytical form. Continuity of this characteristic function and its first two derivatives at the origin of the spectral variable is proven. It is then used to determine the desired conditional mean and conditional variance in a closed analytical form to yield the sequential state estimator. A three-state dynamic system example demonstrates numerically the performance of the Cauchy estimator.**

I. Introduction

In many engineering, economic, telecommunications, and science applications the underlying random processes or noises have significant volatility, which is not captured by the light-tailed Gaussian probability density functions (pdf) [1]. Heavy-tailed distributions have been shown to better represent these volatile random fluctuations. Examples are radar and sonar sensor noise [2] and air turbulent environment noise [3]. Our objective is to develop a filtering technique for linear dynamic systems with heavy-tailed distributed noises while using a particular distribution out of the class of symmetric α-stable distributions [4]. Particular distributions of this class are the Lévy, Gaussian and Cauchy distributions.

The use of heavy tail ($S\alpha$-S) distributions was demonstrated to yield improved filtering and detection results when processing radar signals [5]–[7], radar glint [8], data in a multi-user communication networks [9] and aircraft navigation [10]. These results and the degraded performance of the standard Gaussian estimators when exposed to impulsive noises motivated the derivation of a sequential non-linear estimator for scalar linear dynamic systems with additive Cauchy process and measurement noises [11]. This result was based on propagating the conditional pdf (cpdf) of the system state given the measurement history.

Unfortunately, the recursion scheme for generating the cpdf directly for a scalar linear system [11] does not generalize to the vector state. In [12] a similar estimator was derived by generating the characteristic function of the unnormalized cpdf (ucpdf) in a recursive scheme. This approach is somewhat simpler than the scheme in [11], allows for a stronger result regarding the decay of the estimator parameters with time, and can be generalized to the multi-variable case.

The generalization of the characteristic function of the ucpdf approach [12] to the multi-variable case is the essence of this paper, which is organized as follows. The estimation problem for a n-dimensional, discrete-time, linear system forced by scalar Cauchy process noise and a scalar measurement with additive cauchy measurement noise is formulated in section II. In section III, the characteristic function for the ucpdf of the system state conditioned on the measured history is computed sequentially for the first measurement update and a time propagation step. This motivates the general form of the characteristic function for the ucpdf given in section IV, which is shown to be continuous and twice differentiable in section V and yields closed-form analytical expressions for the minimum variance estimate of the states and the estimation error conditional variance. In section VI a three state system is used to exemplify the performance of the sequential estimator. Concluding remarks are given in section VII.

II. Problem Formulation

We consider the single-input-single-output multivariable linear system

$$x_{k+1} = \Phi x_k + \Gamma w_k, \quad z_k = H x_k + v_k, \tag{1}$$

with state vector $x_k \in \mathrm{R}^n$, scalar measurement z_k, and known matrices $\Phi \in \mathrm{R}^{n\times n}$, $\Gamma \in \mathrm{R}^{n\times 1}$, and $H \in \mathrm{R}^{1\times n}$. The noise inputs are assumed to be independent Cauchy distributed random variables. Specifically, w_k is assumed to be Cauchy distributed with a zero median and a scaling parameter $\beta > 0$. Similarly, v_k has a Cauchy pdf with a median of zero and a scaling parameter $\gamma > 0$. The characteristic functions of these scalar noises are assumed to be time independent and given by

$$\varphi_W(\bar{v}) = e^{-\beta|\bar{v}|}, \quad \varphi_V(\bar{v}) = e^{-\gamma|\bar{v}|}. \tag{2}$$

These characteristic functions have a scalar argument $\bar{v}$. The initial conditions at $k = 1$ are also assumed to be independent Cauchy distributed random variables. Specifically, each i-th element x_{1i} of the initial state vector x_1 has a Cauchy pdf with a zero median and a scaling parameter $\alpha_i > 0$, $i = 1, \ldots, n$. The characteristic function of the joint pdf of the initial conditions, which is a function of a n-dimensional spectral variable $v \in \mathrm{R}^n$, is given by

$$\varphi_{x_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i|v_i|} = \exp\left(-\sum_{i=1}^{n} \alpha_i|v_i|\right) \left[', \exp\left(-\sum_{i=1}^{n} p_i^1|(a_i^1, v)|\right) + j(b_1^1, v)\right]. \tag{3}$$

The last form was introduced for notational convenience to be used in the sequel. We used the definitions

$$p_i^1 = \alpha_i, \quad a_i^1 = e_i, \quad i = 1, \ldots, n, \quad b_1^1 = \{0\}_n, \tag{4}$$

where e_i is a n-dimensional i-th unity vector and $\{0\}_n$ is n-dimensional vector of zeros. With this formulation of the system, our goal is to compute the minimum variance estimate of x_k given the measurement measurement history, i.e., $y_k = z_1 \cdots z_k$.

III. Initial Derivations

The method proposed to solve this Cauchy estimation problem entails propagating the characteristic function of the cpdf of the state vector given a history of measurements. Evaluating this characteristic function and its derivatives at the origin of the spectral vector v will provide the desired state estimate and its error variance. The characteristic function is initiated by the expression given in (3). It changes during a measurement update when a new measurement is processed, and during time propagation affected by the process noise input. We begin by showing how this characteristic function is computed for the first measurement update and time propagation step. This will suggest the general form of the characteristic function and hence the general estimator.

A. Measurement Update at $k = 1$

The characteristic function of the initial state x_1 conditioned on the initial measurement $z_1 = Hx_1 + v_1$ is given by

$$\varphi_{x_1|z_1}(v) = \int_{-\infty}^{\infty} f_{x_1|z_1}(x_1|z_1)\, e^{jv^T x_1} dx_1, \tag{5}$$

where $f_{x_1|z_1}(x_1|z_1) = f_V(z_1 - Hx_1)\, f_{x_1}(x_1) / f_{z_1}(z_1)$ and

$$f_{z_1}(z_1) = \int_{-\infty}^{\infty} f_V(z_1 - Hx_1)\, f_{x_1}(x_1)\, dx_1. \tag{6}$$

The subsequent derivations can be simplified by avoiding the division by $f_{z_1}(z_1)$, thus addressing the unnormalized cpdf and its characteristic function defined by

$$\bar{f}_{x_1|z_1}(x_1|z_1) = f_{x_1|z_1}(x_1|z_1)\, f_{z_1}(z_1) = f_V(z_1 - Hx_1)\, f_{x_1}(x_1), \tag{7}$$

and

$$\bar{\varphi}_{x_1|z_1}(v) = \int_{-\infty}^{\infty} f_V(z_1 - Hx_1)\, f_{x_1}(x_1)\, e^{jv^T x_1} dx_1. \tag{8}$$

When the actual (normalized) functions are needed for, e.g., computing the state estimates, the normalization factor can be easily determined by evaluating $\bar{\varphi}_{x_1|z_1}(x_1|z_1)$ at $v = \{0\}_n$, i.e., $f_{z_1}(z_1) = \bar{\varphi}_{x_1|z_1}(\{0\}_n)$.

Using the derivation in Appendix A, the n integrals in (8) can be expressed as the single integral

$$\bar{\varphi}_{x_1|z_1}(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_{x_1}(v - H^T\eta)\, \varphi_V(-\eta)\, e^{jz_1\eta} d\eta, \tag{9}$$

where $\varphi_V(\cdot)$ and $\varphi_{x_1}(\cdot)$ were defined in (2) and (3), respectively. Hence (9) can be restated as

$$\bar{\varphi}_{x_1|z_1}(v) = \frac{e^{j(b_1^1, v)}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n} p_i^1 |(a_i^1, v) - Ha_i^1 \eta| - \gamma|\eta| + j\,(z_1 - Hb_1^1)\,\eta \right) d\eta. \tag{10}$$

This integral is solved using the general result presented in Appendix B. For that, the coefficients of η in the absolute value term of (10) have to be normalized to one. Clearly, this cannot be attained if some of these coefficients are zero. Hence, we assume that $Ha_i^1 \neq 0$, $i = 1, \ldots, n$, or equivalently that all the elements of H are nonzero. With this assumption, (10) is restated as

$$\bar{\varphi}_{x_1|z_1}(v) = \frac{e^{j(b_1^1, v)}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\sum_{i=1}^{n+1} \rho_i |(\mu_i, v) - \eta| + j\zeta_1\eta \right) d\eta, \tag{11}$$

where we defined

$$\rho_i = p_i^1 |Ha_i^1|, \quad \mu_i = a_i^1/(Ha_i^1), \quad i = 1, \ldots, n,$$
$$\rho_{n+1} = \gamma, \qquad \mu_{n+1} = \{0\}_n, \qquad \zeta_1 = z_1 - Hb_1^1. \tag{12}$$

Using the result in Appendix B, while associating ξ_i and z in the later with (μ_i, v) and ζ_1, respectively, the integral in (11) is evaluated as

$$\bar{\varphi}_{x_1|z_1}(v) = \prod_{i=1}^{n+1} g_i^{1|1}(v_{gi}^{1|1}(v)) \exp\left(y_{ei}^{1|1}(v)\right), \tag{13}$$

where the coefficient functions $g_i^{1|1}(\cdot)$ are given by

$$g_i^{1|1}(y_{gi}^{1|1}(v)) = \frac{1}{2\pi} \left(j\zeta_1 + \rho_i + y_{gi}^{1|1}(v) \right)^{-1} - \left(-j\zeta_1 - \rho_i + y_{gi}^{1|1}(v) \right)^{-1}. \tag{14}$$

The arguments of $g_i^{1|1}(\cdot)$ and the exponents in (13) are

$$y_{gi}^{1|1}(v) = \sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l \operatorname{sgn}((a_l, v)), \tag{15a}$$

$$y_{ei}^{1|1}(v) = -\sum_{\substack{l=1 \\ l \neq i}}^{n+1} \rho_l |(a_l, v)| + j(b_i^{1|1}, v), \tag{15b}$$

where, while using (12), for $R \neq i$ we have defined

$$a_i = \begin{cases} \dfrac{a_1^1}{Ha_1^1} - \dfrac{a_i^1}{Ha_i^1} & i \neq n+1,\ R \neq n+1 \\ Ha_1^1 & i = n+1 \end{cases} \tag{16a}$$

$$b_i^{1|1} = \begin{cases} \dfrac{a_1^1}{Ha_1^1} - Ha_i^1 & R = n+1 \\ \left(z_1 - Hb_1^1\right) \dfrac{a_i^1}{Ha_i^1} + b_1^1 & i \neq n+1 \\ b_1^1 & i = n+1 \end{cases} \tag{16b}$$

To simplify the notation and subsequent derivations, specifically to avoid the $R \neq i$ exclusion in the sums of (15), the elements in these sums are renumbered sequentially, i.e., $R \in [1, n]$ for each i-th term. The renumbered vectors a_i will be denoted by $a_i^{1|1}$. To accommodate subsequent derivations, the coefficients ρ in the sums of (15) will be marked differently for $y_{gi}^{1|1}(\cdot)$ and $y_{ei}^{1|1}(\cdot)$: we will use $q_i^{1|1}$ for the former and $p_i^{1|1}$ for the latter. The parameters z and ρ_i in (14) will be denoted by $c_i^{1|1}$ and $d_i^{1|1}$, respectively. Finally, the number of terms in the sum of (13) will be denoted by $n_t^{1|1} = n+1$, while the number of elements in the sums of (15) will be marked by $n_{ei}^{1|1} = n$. Although at this stage all the counters $n_{ei}^{1|1}$ are the same, as will be seen in the sequel, they may be different for different i-s. Therefore, we have introduced i-dependent element counters $n_{ei}^{1|1}$. Moreover, although the counters $n_t^{1|1}$ and $n_{ei}^{1|1}$ seem to be related to each other (the former is $n+1$ while the latter are n), they will exhibit irregular changes and thus were introduced separately.

With these notations and renumbered parameters and vectors, $\bar{\varphi}_{X_1|Z_1}(v)$ is restated as

$$\bar{\varphi}_{X_1|Z_1}(v) = \sum_{i=1}^{n_t^{1|1}} g_i^{1|1}\left(y_{gi}^{1|1}(v)\right) \exp\left(y_{ei}^{1|1}(v)\right), \qquad (17)$$

where

$$g_i^{1|1}\left(y_{gi}^{1|1}(v)\right) = \frac{1}{2\pi}\left(jc_i^{1|1} + d_i^{1|1} + y_{gi}^{1|1}(v)\right)^{-1} - \left(-jc_i^{1|1} - d_i^{1|1} + y_{gi}^{1|1}(v)\right)^{-1} \qquad (18)$$

and

$$v_{gi}^{1|1}(v) = \sum_{i=1}^{n_{ei}^{1|1}} q_i^{1|1}\operatorname{sgn}\left(a_i^{1|1}, v\right), \qquad (19a)$$

$$y_{ei}^{1|1}(v) = -\sum_{i=1}^{n_{ei}^{1|1}} p_i^{1|1}\left|\left(a_i^{1|1}, v\right)\right| + j\left(b_i^{1|1}, v\right). \qquad (19b)$$

It is interesting to note that the initial characteristic function in (3) is also expressed in a form identical to that given in (17). The initial parameters are: $n^1 = 1$; $g^1 = 1$ (and thus there are no y_{g1}^1 arguments); $n_e^1 = n$; the parameters p_i^1 and the vectors a_i^1 and b_i^1 are given in (4).

Remark 3.1: The special cases when some of the conditions $Ha_i^1 \neq 0$, $i = 1, \ldots, n$ do not hold can be handled similarly, to yield results that resemble those presented above. These cases are not presented here for brevity.

B. Time Propagation to $k = 2$

Given the time propagation equation $x_2 = \Phi x_1 + \Gamma w_1$, while the ucpdf of x_2 given z_1 is given by

$$\bar{\varphi}_{X_2|Z_1}(v) = \bar{\varphi}_{X_1|Z_1}(\Phi^T v)\,\varphi_W(\Gamma^T v) = \sum_{i=1}^{n_t^{1|1}} g_i^{1|1}\left(y_{gi}^{1|1}(\Phi^T v)\right) \exp\left(y_{ei}^{1|1}(\Phi^T v) - \beta\left|(\Gamma, v)\right|\right). \qquad (20)$$

The last expression in (20) can be expressed in a form that is similar to the one given in (17). For that, we make the following observations that lead to the definitions of new and time propagated parameters and vectors.

1) $\bar{\varphi}_{X_2|Z_1}(v)$ is expressed with the same number of terms as $\bar{\varphi}_{X_1|Z_1}(v)$. Hence, we define $n_t^{2|1} = n_t^{1|1}$.
2) The coefficient functions $g_i^{1|1}(\cdot)$, or specifically the parameters $c_i^{1|1}$ and $d_i^{1|1}$, remain as in (18), i.e., $c_i^{2|1} = c_i^{1|1}$, $d_i^{2|1} = d_i^{1|1}$, $i \in [1, \ldots, n_t^{1|1}]$. There are only changes in the arguments of $g_i^{1|1}(\cdot)$. For notational consistency we will denote the updated coefficient functions $g_i^{2|1}(\cdot)$ while remembering that
$$g_i^{2|1}(\cdot) = g_i^{1|1}(\cdot). \qquad (21)$$
3) The parameters $q_i^{1|1}$ and $p_i^{1|1}$ used to define $y_{gi}^{1|1}(\cdot)$ and $y_{ei}^{1|1}(\cdot)$ in (19) are also unchanged, i.e., $p_i^{2|1} = p_i^{1|1}$, $q_i^{2|1} = q_i^{1|1}$, $i \in [1, \ldots, n_t^{2|1}]$, $R \in [1, \ldots, n_{ei}^{1|1}]$.
4) The arguments of $y_{gi}^{1|1}(\cdot)$ and $y_{ei}^{1|1}(\cdot)$ are multiplied by Φ^T, hence affecting the vectors $a_i^{1|1}$ and $b_i^{1|1}$ as
$$y_{gi}^{1|1}(\Phi^T v) = \sum_{i=1}^{n_{ei}^{1|1}} q_i^{2|1}\operatorname{sgn}\left(a_i^{2|1}, v\right), \qquad (22a)$$
$$v_{ei}^{1|1}(\Phi^T v) = -\sum_{i=1}^{n_{ei}^{1|1}} p_i^{2|1}\left|\left(a_i^{2|1}, v\right)\right| + j\left(b_i^{2|1}, v\right), \qquad (22b)$$
where we used the definitions $a_i^{2|1} = \Phi a_i^{1|1}$, $b_i^{2|1} = \Phi b_i^{1|1}$, $i \in [1, \ldots, n_t^{2|1}]$, $R \in [1, \ldots, n_{ei}^{1|1}]$.
5) The exponents are a function of an additional element $-\beta\left|(\Gamma, v)\right|$. Hence, the number of elements that define the new argument $y_{ei}^{2|1}(\cdot)$ increase by one, i.e., $n_{ei}^{2|1} = n_{ei}^{1|1} + 1$. The parameters and vectors that define these new elements are $p_{in_{ei}^{2|1}}^{2|1} = \beta$, $a_{in_{ei}^{2|1}}^{2|1} = \Gamma$, $i \in 1, \ldots, n_t^{2|1}]$. For consistency, and to facilitate the subsequent manipulations of the characteristic function, the number of elements in the sum of the new arguments $y_{gi}^{2|1}(\cdot)$ are also increased by one, while introducing the zero parameters $q_{in_{ei}^{2|1}}^{2|1} = 0$, $i \in [1, \ldots, n_t^{2|1}]$. Hence, with these new elements, the arguments $y_{gi}^{2|1}(\cdot)$ and $v_{ei}^{2|1}(\cdot)$ are defined as
$$v_{gi}^{2|1}(v) = \sum_{i=1}^{n_{ei}^{2|1}} q_i^{2|1}\operatorname{sgn}\left(a_i^{2|1}, v\right) \qquad (23a)$$

using the result of Appendix C, the

characteristic function of

$$y_{ei}(v) = -\sum_{=1} p_i \ |(a_i \ , v)| + j(b_i \ , v). \qquad (23b)$$

Using all the time propagated and newly defined parameters, (20) is restated as

$$\bar{\varphi}_{X_2|Z_1}(v) = \sum_{i=1}^{n_t^{2|1}} g_i^{2|1}\left(y_{gi}^{2|1}(v)\right) \exp\left(y_{ei}^{2|1}(v)\right), \quad (24)$$

where

$$g_i^{2|1}\left(y_{gi}^{2|1}(v)\right) = \frac{1}{2\pi}\left(jc_i^{2|1} + d_i^{2|1} + y_{gi}^{2|1}(v)\right)^{-1}\left(-jc_i^{2|1} - d_i^{2|1} + y_{gi}^{2|1}(v)\right)^{-1} \quad (25)$$

with arguments $y_{gi}^{2|1}(\cdot)$ and $y_{ei}^{2|1}(\cdot)$ given in (23). Overall, we have obtained a form which is similar to the one in (17), determined after the first measurement update in the previous subsection.

C. Summary of Initial Results

The above derivations demonstrate that the characteristic function of the ucpdf of the state x_k at time steps $k = 1$ and 2 is expressed as a sum of n_t weighted exponential terms. The number of terms increases during a measurement update and are unchanged during a time propagation step. In each such term i, the exponents and their weights or coefficients are functions of a sum of n_{ei} elements. The number n_{ei} of elements normally increases during the time propagation step and are unchanged during a measurement update. Those observations provide the insight and guidance on how to construct the characteristic functions at any time step k, as is discussed next.

IV. Measurement Update and Time Propagation: General Case

The initial results presented in the previous section suggest the general form of the characteristic function of the ucpdf of the state at any time step k given a measurement history. Consequently, we assume that at any time step the characteristic function of the ucpdf of the state x_k given all the data history up to k, i.e., $y_k = z_1 \cdots z_k$, is expressed as

$$\bar{\varphi}_{X_k|Y_k}(v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(v)\right) \exp\left(y_{ei}^{k|k}(v)\right), \quad (26)$$

where

$$g_i^{k|k}\left(y_{gi}^{k|k}(v)\right) = \frac{1}{2\pi}\left(jc_i^{k|k} + d_i^{k|k} + y_{gi2}^{k|k}(v)\right)^{-r_i^{k|k}} g^{k-1|k-1}\left(y_{gi1}^{k|k}(v) + h_i^{k|k}\right) \left(-jc_i^{k|k} - d_i^{k|k} + y_{gi2}^{k|k}(v)\right)^{-r_i^{k|k}} g^{k-1|k-1}\left(y_{gi1}^{k|k}(v) - h_i^{k|k}\right), \quad (27)$$

and

$$y_{gi}^{k|k}(v) = \sum_{l=1}^{n_{ei}^{k|k}} q_{il}^{k|k} \operatorname{sgn}\left(a_{il}^{k|k}, v\right) \in \mathbb{R}^k, \quad (28a)$$

$$y_{ei}^{k|k}(v) = -\sum_{l=1}^{n_{ei}^{k|k}} p_{il}^{k|k} \left|\left(a_{il}^{k|k}, v\right)\right| + j\left(b_{il}^{k|k}, v\right). \quad (28b)$$

As will be explained below, $y_{gi}^{k|k}(\cdot)$ and $q_{il}^{k|k}$ are k dimensional vectors. When evaluating $g_i^{k|k}\left(y_{gi}^{k|k}(\cdot)\right)$ in (27), the argument $y_{gi}^{k|k}(\cdot)$ is partitioned as follows: $y_{gi1}^{k|k}(\cdot)$ is a $k-1$ dimensional vector constructed from its first $k-1$ components, while the scalar $y_{gi2}^{k|k}(\cdot)$ is its last component. Also, the source of the indexes l and offsets h_i will be detailed in the measurement update step. The initialization of this function at $k = 1$ is performed based on the results derived in subsection III-A. Now we will perform one time propagation and one measurement update and show that the above form is maintained at any time step.

A. Time Propagation from k to $k+1$

The characteristic function $\bar{\varphi}_{X_{k+1}|Y_k}(v)$ of the time propagated state $x_{k+1} = \Phi x_k + \Gamma w_k$ is determined using the result in (C.3) of Appendix C and is expressed as

$$\bar{\varphi}_{X_{k+1}|Y_k}(v) = \bar{\varphi}_{X_k|Y_k}(\Phi^T v)\,\varphi_W(\Gamma^T v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\Phi^T v)\right) \exp\left(y_{ei}^{k|k}(\Phi^T v) - \beta\left|(\Gamma, v)\right|\right). \quad (29)$$

The arguments of the coefficient functions $g_i^{k|k}(\cdot)$ above and those of the exponents in (29) are redefined as

$$y_{gi}^{k+1|k}(v) = \sum_{l=1}^{n_{ei}^{k|k}} q_{il}^{k|k} \operatorname{sgn}\left(\Phi a_{il}^{k|k}, v\right), \quad (30a)$$

$$y_{ei}^{k+1|k}(v) = -\sum_{l=1}^{n_{ei}^{k|k}} p_{il}^{k|k} \left|\left(\Phi a_{il}^{k|k}, v\right)\right| - \beta\left|(\Gamma, v)\right| + j\left(b_{il}^{k|k}, v\right). \quad (30b)$$

With these definitions, (29) can be restated as

$$\bar{\varphi}_{X_{k+1}|Y_k}(v) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k+1|k}(v)\right) \exp\left(y_{ei}^{k+1|k}(v)\right), \quad (31)$$

which clearly has a form similar to (26). Specifically, defining the time propagated parameters similarly to the five-steps procedure outlined in section III-B including the new parameters $r_i^{k+1|k} = r_i^{k|k}$, $h_i^{k+1|k} = h_i^{k|k}$, the result above can be restated as

$$\bar{\varphi}_{X_{k+1}|Z_k}(v) = \sum_{i=1}^{n_t^{k+1|k}} g_i^{k+1|k}\left(y_{gi}^{k+1|k}(v)\right) \exp\left(y_{ei}^{k+1|k}(v)\right), \quad (32)$$

where

$$g_i^{k+1|k}\left(y_{gi}^{k+1|k}(v)\right) = \frac{1}{2\pi}\int \frac{g_{r_i^{k+1/k}}^{k-1|k-1}\left(y_{gi1}^{k+1|k}(v)+h_i^{k+1|k}\right)}{jc_i^{k+1|k}+d_i^{k+1|k}+y_{gi2}^{k+1|k}(v)} - \frac{g_{r_i^{k+1/k}}^{k-1|k-1}\left(y_{gi1}^{k+1|k}(v)-h_i^{k+1|k}\right)}{jc_i^{k+1|k}-d_i^{k+1|k}+y_{gi2}^{k+1|k}(v)} \quad (33)$$

and

$$v_{gi}^{k+1|k}(v) = \sum_{l=1}^{n_{ei}^{k+1/k}} q_{il}^{k+1|k}\,\mathrm{sgn}\left(a_{il}^{k+1|k}, v\right) \in \mathbb{R}^k, \quad (34a)$$

$$y_{ei}^{k+1|k}(v) = -\sum_{l=1}^{n_{ei}^{k+1/k}} p_{il}^{k+1|k}\left|\left(a_{il}^{k+1|k}, v\right)\right| + j\left(b_{il}^{k+1|k}, v\right). \quad (34b)$$

The number of terms that define $\bar{\varphi}_{X_{k+1}|Z_k}(\cdot)$ in (32) is identical to that of $\bar{\varphi}_{X_k|Z_k}(\cdot)$ in (26). However, the number of elements needed to define the arguments of those terms has increased by one, i.e., $n_{ei}^{k+1|k} = n_{ei}^{k|k} + 1$.

B. Measurement Update at $k+1$

The measurement $z_{k+1} = Hx_{k+1} + v_{k+1}$ is processed next to determine $\bar{\varphi}_{X_{k+1}|Y_{k+1}}(\cdot)$, where $y_{k+1} = \begin{bmatrix} z_1 & \cdots & z_{k+1} \end{bmatrix}$. Using the general result in Appendix A,

$$\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}_{X_{k+1}|Y_k}\left(v - H^T\eta\right)\exp\left(-\gamma|\eta| + jz_{k+1}\eta\right)d\eta, \quad (35)$$

where $\bar{\varphi}_{X_{k+1}|Z_k}(\cdot)$ was determined in (32). This integral was solved in (B.2) of Appendix B. This result indicates that each term in the sum of $\bar{\varphi}_{X_{k+1}|Z_k}(\cdot)$ in (32) generates $n_{ei}^{k+1|k}+1$ terms in $\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v)$: the index $r_i^{k+1|k+1}$ indicates which term of $\bar{\varphi}_{X_{k+1}|Z_k}(\cdot)$ generated the new ones. In addition, this parent term is called with an offset, denoted as Q_i in (B.2). In the final result, this offset will be denoted as $h_i^{k+1|k+1}$. Finally we note that the coefficient function of the exponents in (B.2) have two input arguments: the sum in the numerators (with the offset) and the sum in the denominators. Grouping those two arguments into one, we denote the input vector to the updated coefficient function $g_i^{k+1|k+1}(\cdot)$ as $y_{ei}^{k+1|k+1}(\cdot)$. Clearly, each measurement update increases the dimension of this argument, and hence the deduction in (28) that $y_{gi}^{k|k}(\cdot)$ and $q_i^{k|k}$ are k dimensional vectors. Similarly, the updated $y_{gi}^{k+1|k+1}(\cdot)$ and $q_i^{k+1|k+1}$ will be of dimension $k+1$.

After the integration, reordering terms and defining updated parameters similar to the procedure outlined in III-A, the measurement updated characteristic function is expressed as

$$\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v) = \sum_{i=1}^{n_t^{k+1/k+1}} g_i^{k+1|k+1}\left(y_{gi}^{k+1|k+1}(v)\right) \times \exp\left(y_{ei}^{k+1|k+1}(v)\right), \quad (36)$$

$$g_i^{k+1|k+1}\left(y_{gi}^{k+1|k+1}(v)\right) = \frac{1}{2\pi}\int \frac{g_{r_i^{k+1/k+1}}^{k|k}\left(y_{gi1}^{k+1|k+1}(v)+h_i^{k+1|k+1}\right)}{jc_i^{k+1|k+1}+d_i^{k+1|k+1}+y_{gi2}^{k+1|k+1}(v)} - \frac{g_{r_i^{k+1/k+1}}^{k|k}\left(y_{gi1}^{k+1|k+1}(v)-h_i^{k+1|k+1}\right)}{jc_i^{k+1|k+1}-d_i^{k+1|k+1}+y_{gi2}^{k+1|k+1}(v)}, \quad (37)$$

and

$$y_{gi}^{k+1|k+1}(v) = \sum_{l=1}^{n_{ei}^{k+1/k+1}} q_{il}^{k+1|k+1}\,\mathrm{sgn}\left(a_{il}^{k+1|k+1}, v\right), \quad (38a)$$

$$y_{ei}^{k+1|k+1}(v) = -\sum_{l=1}^{n_{ei}^{k+1/k+1}} p_{il}^{k+1|k+1}\left|\left(a_{il}^{k+1|k+1}, v\right)\right| + j\left(b_{il}^{k+1|k+1}, v\right). \quad (38b)$$

The above results clearly show that the form of the characteristic function proposed in (26-28) for the time step k is also maintained at $k+1$.

V. Conditional Mean and Estimation Error Variance

The minimum conditional variance estimator of x_k given the measurement sequence $y_k = \begin{bmatrix} z_1 & \cdots & z_k \end{bmatrix}$ is the conditional mean of x_k given y_k. It can be determined by evaluating the characteristic function of (26) and its derivatives at $v = \{0\}_n$, or as $v \to \{0\}_n$. In this section we show that $\bar{\varphi}_{X_k|Y_k}(\cdot)$ is twice continuously differentiable and give explicit expressions for the conditional mean and the estimation error–variance.

The continuity of the first two derivatives of $\bar{\varphi}(\cdot)$ is proven by induction. The characteristic function $\varphi_{X_k|Y_k}(v)$ is given by the convolution integral in (11). Using the definitions in (4) and (12), (11) is rewritten as

$$\bar{\varphi}_{X_1|Z_1}(\tilde{v}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\prod_{i=1}^{n} e^{\alpha_i|\tilde{v}_i-\eta|}e^{-\gamma|\eta|+jz_1\eta}d\eta \quad (39)$$

where $\tilde{v}_i \triangleq (\mu_i, v)$ and $\tilde{v} = \begin{bmatrix} \tilde{v}_1 & \cdots & \tilde{v}_n \end{bmatrix}^T$. The integrand of (39) is a continuous function of η and $\tilde{v}_i \forall i = 1, \ldots, n$. Moreover, it can be shown that its first and, by a change of variables, its second order partial derivatives with respect to any $\tilde{v}_j$ are piecewise continuous and bounded. Consequently, $\varphi_{X_1|Z_1}(\tilde{v})$ is twice continuously differentiable with respect to $\tilde{v}$ [13]. Since v is a linear function of $\tilde{v}$, $\bar{\varphi}_{X_1|Z_1}(v)$ is twice continuously differentiable with respect to v.

The time propagated characteristic function at $k = 2$ of (20) is restated as

$$\bar{\varphi}_{X_2|Z_1}(v) = \bar{\varphi}_{X_1|Z_1}(\Phi^T v)\, e^{-\beta|\Gamma^T v|}. \quad (40)$$

Since, as it was shown above, $\bar{\varphi}_{X_1|Z_1}(v)$ is twice continuously differentiable, $\bar{\varphi}_{X_1|Z_1}(\Phi^T v)$ is also twice continuously differentiable for any transition matrix Φ. Clearly, $\bar{\varphi}_{X_2|Z_1}(v)$ is continuous, being a product of two continuous functions. However, since the first derivative of $e^{-\beta|\Gamma^T v|}$ is not continuous at $v = \{0\}_n$, the associated pdf $f_{X_2|Z_1}(x_2|z_1)$ does not have any moments. This implies that we cannot compute *a priori* estimates of the state x_2 given only the past measurement z_1.

The characteristic function at the second measurement update given the measurement history $y_2 = \{z_1, z_2\}$, while using the explicit form of (40) is given by

$$\bar{\varphi}_{X_2|Y_2}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{\varphi}_{X_2|Z_1}(v - H^T\eta)\, e^{-\beta|\Gamma^T v - H\Gamma\eta|} e^{-\gamma|\eta| + jz_2\eta}\, d\eta, \quad (41)$$

where $\tilde{\varphi}_{X_2|Z_1}(v) \triangleq \bar{\varphi}_{X_1|Z_1}(\Phi^T v)$. Note that if $H\Gamma = 0$, the term $e^{-\beta|\Gamma^T v|}$ would come out of the integral in (41). Consequently, in this case $\varphi_{X_2|Z_2}(v)$ would not be continuously differentiable with respect to v, and there would be no minimum variance estimate of x_k given y_k for all $k \geq 2$. Therefore, $H\Gamma \neq 0$ is a necessary condition for the continuous differentiability of $\bar{\varphi}_{X_2|Y_2}(v)$ and the existence of the desired estimate of the state x_k at any time step k.

It was shown earlier that $\tilde{\varphi}_{X_2|Z_1}(v - H^T\eta)$ is twice continuously differentiable with respect to v. Moreover, $e^{-\beta|\Gamma^T v - H\Gamma\eta|}$ has a piecewise continuous and bounded first derivative with respect to v. Hence, assuming $H\Gamma \neq 0$, the integrand in (41) can be shown to have piecewise continuous and bounded first and second order partial derivatives with respect to any v. This implies that $\bar{\varphi}_{X_2|Z_2}(v)$ is twice continuously differentiable with respect to v [13].

In a recursive manner, the continuity of the first two derivatives of the characteristic function can be shown to be maintained when propagating from any time step k to time step $k+1$, as it was maintained in going from times step 1 to 2 presented above. Assuming that $\bar{\varphi}_{X_k|Y_k}(v)$ of (26) is twice continuously differentiable, one could use the above arguments to show that so is $\bar{\varphi}_{X_{k+1}|Y_{k+1}}(v)$.

Having established that $\bar{\varphi}_{X_k|Y_k}(v)$ is twice continuously differentiable, the explicit form of the pdf of the measurement history, the conditional mean, and conditional variance is determined from (26-28) by evaluating $\bar{\varphi}_{X_k|Y_k}(v)$ and its derivative as $v \to \{0\}_n$. We choose $v = \hat{v}$, where ρ is a positive scalar such that $\rho \to 0$ and $\hat{v}$ is any fixed direction in the v domain for which $(a_i^{k|k}, \hat{v}) \neq 0, \ \forall\ (i, R)$. This condition avoids the discontinuity issues of $\mathrm{sgn}(a_i^{k|k}, \hat{v})$ when evaluating $\bar{\varphi}_{X_k|Y_k}(v)$ and its derivatives. With this choice of $\hat{v}$, $\mathrm{sgn}(a_i^{k|k}, \hat{v}) = \mathrm{sgn}(a_i^{k|k}, v)$ [7], s^i are well defined constants that satisfy $|s^i| = 1$. From (26-28), $\bar{\varphi}_{X_k|Y_k}(\cdot)$ along this direction is given by

$$\bar{\varphi}_{X_k|Y_k}(\rho\hat{v}) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\hat{v})\right) \exp\left(y_{ei}^{k|k}(\hat{v})\right), \quad (42)$$

where the functions $g_i^{k|k}(\cdot)$ are given in (27). Their arguments, defined in (28), are

$$y_{gi}^{k|k}(\hat{v}) = \sum_{=1}^{n_{ei}^{k|k}} q_i\, s^{=} = y_{gi}^{k|k}(\hat{v}). \quad (43)$$

Since the vector $\hat{v}$ is constant, the arguments $y_{gi}^{k|k}(\cdot)$ and thus the coefficient functions $g_i^{k|k}(\cdot)$ are constant along the chosen direction.

Similarly, the arguments $y_{ei}^{k|k}(\cdot)$ of the exponents in (42), defined in (28), are manipulated as follows

$$y_{ei}^{k|k}(\hat{v}) = -\sum_{=1}^{n_{ei}^{k|k}} p_i^{k|k}\,|(a_i^{k|k}, \hat{v})| + j(b_i^{k|k}, \hat{v}) \quad (44)$$

$$= -\sum_{=1}^{n_{ei}^{k|k}} p_i^{k|k}\, s^i\,(a_i^{k|k}, \hat{v}) + j(b_i^{k|k}, \hat{v}) = (\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v}),$$

where we have defined

$$\bar{y}_{ei}^{k|k}(\hat{v}) = -\sum_{=1}^{n_{ei}^{k|k}} p_i^{k|k} s^i a_i^{k|k} + j b_i^{k|k}. \quad (45)$$

Since $\hat{v}$ is constant, so is the expression $(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})$, making $y_{ei}^{k|k}(\cdot)$ of the exponents linear in ρ along the chosen direction. Using the above results, (42) can be restated as

$$\bar{\varphi}_{X_k|Y_k}(\rho\hat{v}) = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\hat{v})\right) \exp\left((\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})\,\rho\right). \quad (46)$$

The conditional mean of the state x_k given the data sequence y_k is given by

$$\hat{x}_k = E[x_k|y_k] = \frac{1}{j f_{Y_k}(y_k)} \left(\frac{\partial \bar{\varphi}_{X_k|Y_k}(\rho\hat{v})}{\partial(\rho\hat{v})}\right)^T. \quad (47)$$

The pdf $f_{Y_k}(y_k)$ needed to normalize the above result is determined by evaluating $\bar{\varphi}_{X_k|Y_k}(\rho\hat{v})$ at $\rho = 0$. Using (46)

$$f_{Y_k}(y_k) = \bar{\varphi}_{X_k|Y_k}(\rho\hat{v})\Big|_{\rho=0} = \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\hat{v})\right). \quad (48)$$

The derivative of $\bar{\varphi}_{X_k|Y_k}(\cdot)$ is determined by differentiating (46). Since $g_i^{k|k}(y_{gi}^{k|k}(\hat{v}))$ and $(\bar{y}_{ei}^{k|k}(\hat{v}), \hat{v})$ are constant, this derivative can be easily determined. Evaluating it at $\rho = 0$ yields the minimum conditional variance estimate

$$\hat{x}_k = \frac{1}{j f_{Y_k}(y_k)} \sum_{i=1}^{n_t^{k|k}} g_i^{k|k}\left(y_{gi}^{k|k}(\hat{v})\right)\, \bar{y}_{ei}^{k|k}(\hat{v}). \quad (49)$$

The second conditional moment of x_k given y_k is determined by

$$E[x_k x_k^T | y_k] = \frac{1}{j^2 f_{Y_k}(y_k)} \left. \frac{\partial^2 \bar{\varphi}_{x_k|Y_k}(\hat{v})}{\partial(\hat{v})\partial(\hat{v})^T} \right|_{=0} . \quad (50)$$

The second derivative in the above equation is easily computed by twice differentiating (46) to yield

$$E[x_k x_k^T | y_k] = \quad (51)$$
$$\frac{1}{j^2 f_{Y_k}(y_k)} \sum_{i=1}^{n_{k|k}^{k|k}} g_{k|k}^{i} \left(y_{gi}^{k|k}(\hat{v}) \right) \left(\bar{y}_{ei}^{k|k}(\hat{v}) \right) \left(\bar{y}_{ei}^{k|k}(\hat{v}) \right)^T .$$

Finally, from (49) and (51), the variance of the estimation error $\tilde{x}_k = x_k - \hat{x}_k$ can be evaluated as $E[\tilde{x}_k \tilde{x}_k^T | y_k] = E[x_k x_k^T | y_k] - \hat{x}_k \hat{x}_k^T$.

VI. Three-Dimensional State Estimator

The performance of the proposed estimator was tested numerically. The system parameters were chosen as follows:

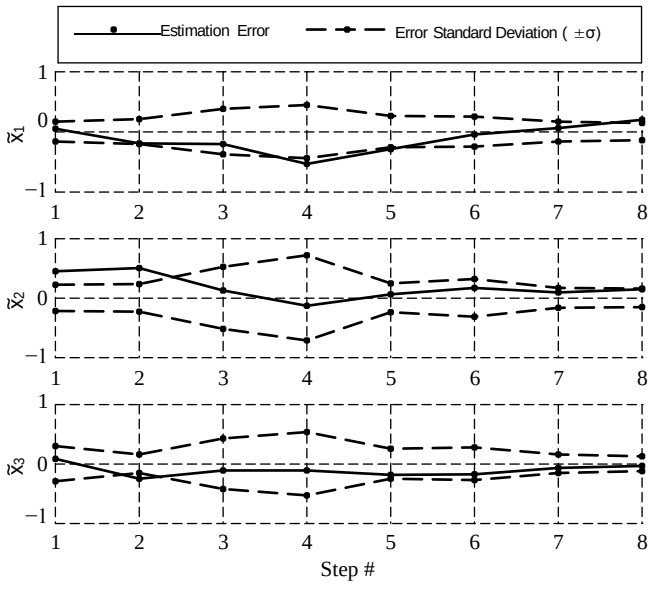
$$\Phi = \begin{bmatrix} 1.4 & -0.6 & -1.0 \\ -0.2 & 1.0 & 0.5 \\ 0.6 & -0.6 & -0.2 \end{bmatrix}, \Gamma = \begin{bmatrix} 0.1 & 0 & 0 \\ 0.3 & \alpha_1 & 0 \\ -0.2 & 0 & \alpha_3 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0.1 \\ 0.08 \\ 0.05 \end{bmatrix},$$
$$H = \begin{bmatrix} 1.0 & 0.5 & 0.2 \end{bmatrix}, \beta = 0.1, \quad \gamma = 0.2.$$

The system has stable eigenvalues at $0.7 \pm 0.3j$ and 0.8. It is observable and complies with the necessary condition that $H\Gamma \neq 0$.

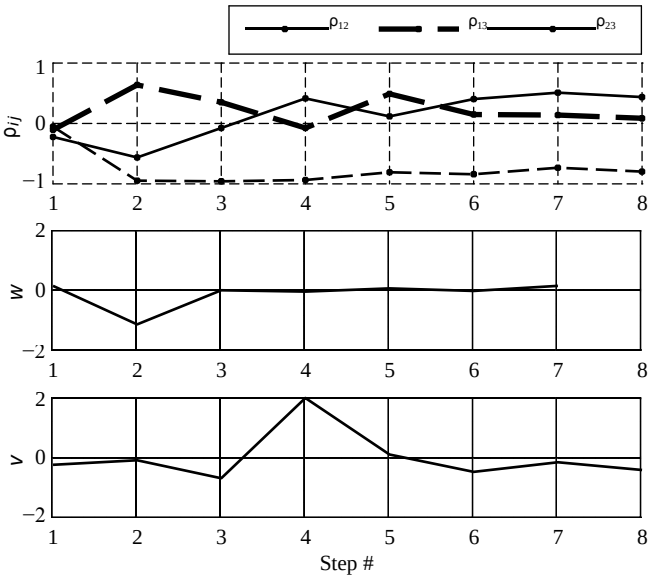
The simulation results are depicted in Fig. 1, showing the estimation errors in 1(a), and the cross-correlation, process and measurement noises in 1(b). It can be clearly observed that the estimation error standard deviation depends on the measurement sequence. Specifically, the standard deviation increases when a large measurement noise is encountered, e.g., at time step #4.

VII. Conclusions

For a linear discrete-time system with additive Cauchy measurement and process noises, an analytic recursion of the characteristic function of the unnormalized condition pdf was determined for the n-vector state estimator. It was then shown that this characteristic function is twice continuously differentiable, allowing the determination of the conditional mean and the conditional second moment, from which the conditional error variance is determined. This conditional mean constitutes the Cauchy estimator for the multivariable system. The estimator was then evaluated numerically for a third-order example.



(a) Estimation Errors



(b) Cross-correlations and system noises

Fig. 1. Three-State Estimation Example: Numerical Results

References

[1] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

[2] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution," *IEEE Communications Letters*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[3] P. Reeves, "A non-gaussian turbulence simulation," Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-69-67, 1969.

[4] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall, 1994.

[5] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[6] G. R. Arce, *Nonlinear Signal Processing: A statistical Approach*. Wiley, New Jersey, 2005.

[7] J. P. Nolan, J. G. Gonzalez, and N. nez, "Stable filters: A robust signal processing framework for heavy-tailed noise." in *In Proceedings of the 2010 IEEE Radar Conference*, 2010.

[8] G. A. Hewer, R. D. Martin, and J. Zeh, "Robust preprocessing for kalman filtering of glint noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 120–128, January 1987.

[9] G. A. Tsihrintzis, *Statistical Modeling and Receiver Design for Multi-*

User Communication Networks; in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.
[10] C. J. Masreliez and R. D. Martin, "Robust bayesian estimation for the linear model and robustifying the kalman filter," *IEEE Transactions on Automatic Control*, vol. AC-22, no. 3, pp. 361–371, June 1977.
[11] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1329–1342, 2010.
[12] ——, "State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach," *SIAM J. Control Optim.*, vol. 0, no. 0, pp. 1–24, 2012.
[13] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*. Wiley-Interscience, 1984.
[14] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

APPENDIX

A. Characteristic Function of an Unnormalized Conditional PDF Given a Scalar Measurement

For the linear measurement equation $z = Hx + v = \sum_{i=1}^{n} h_i x_i + v$, the characteristic function of the unnormalized conditional density function of the state is given as

$$\bar{\varphi}_{x|z}(v) = \int_{-\infty}^{\infty} f_X(x)\, f_V(z - Hx)\, e^{jv^T x} dx. \quad \text{(A.1)}$$

The above resembles a Fourier transform of a product of two functions: $f_X(x)$ and $f_V(z - Hx)$. Using the dual convolution property, this integral can be solved by a convolution in the v domain between the characteristic function $\varphi_X(v)$ of $f_X(x)$ and the characteristic function of $f_V(z - Hx)$, which we denote by $\hat{\varphi}_V(v)$, i.e.

$$\bar{\varphi}_{x|z}(v) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \varphi_X(v - \sigma)\, \hat{\varphi}_V(\sigma)\, d\sigma. \quad \text{(A.2)}$$

The characteristic function $\hat{\varphi}_V(v)$ is determined as follows

$$\hat{\varphi}_V(v) = \int_{-\infty}^{\infty} f_V\left(z - \sum_{i=1}^{n} h_i x_i\right) e^{j \sum_{i=1}^{n} v_i x_i} dx. \quad \text{(A.3)}$$

Substituting (A.3) into (A.2) and carrying out the resulting $n - 1$ nested integrations involving $n - 1$ delta functions, it is straight forward to show that

$$\bar{\varphi}_{x|z}(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(v - H^T\eta)\, \varphi_V(-\eta)\, e^{jz\eta} d\eta. \quad \text{(A.4)}$$

B. Integral of an Exponent of Absolute Values

The measurement update stage of the Cauchy estimator entails evaluating a convolution integral of the following form

$$I = \int_{-\infty}^{\infty} g\left(\sum_{i=1}^{n} Q_i \operatorname{sgn}(\xi_i - \eta)\right) \times \exp\left(-\sum_{i=1}^{n} \rho_i |\xi_i - \eta| + jz\eta\right) d\eta, \quad \text{(B.1)}$$

where Q_i could be m-dimensional vectors. Assuming that the constants ξ_i are ordered such that $\xi_i < \xi_{i+1}$, the integral in (B.1) can be solved over the intervals $\xi_l \le \eta < \xi_{l+1}$ in which the $\operatorname{sgn}(\cdot)$ functions are constants equal to ± 1 and the absolute value can be removed from the exponents. Summing those integrals yields the closed form result

$$I = \sum_{l=1}^{n} \exp\left(-\sum_{i=1}^{n} \rho_i |\xi_i - \xi_l| + jz\xi_l\right) \times g\left(Q_l + \sum_{\substack{i=1 \\ i \neq l}}^{n} Q_i \operatorname{sgn}(\xi_l - \xi_i)\right) \times \left(jz + \rho_l + \sum_{\substack{i=1 \\ i \neq l}}^{n} \rho_i \operatorname{sgn}(\xi_l - \xi_i)\right)^{-1} - \sum_{l=1}^{n} \exp\left(-\sum_{i=1}^{n} \rho_i |\xi_i - \xi_{l-1}| + jz\xi_{l-1}\right) \times g\left(-Q_{l-1} + \sum_{\substack{i=1 \\ i \neq l-1}}^{n} Q_i \operatorname{sgn}(\xi_{l-1} - \xi_i)\right) \times \left(jz - \rho_{l-1} + \sum_{\substack{i=1 \\ i \neq l-1}}^{n} \rho_i \operatorname{sgn}(\xi_{l-1} - \xi_i)\right)^{-1}. \quad \text{(B.2)}$$

C. Time Propagation of Characteristic Functions for Linear Systems

Assume that at step k the state pdf and characteristic function are given by $f_{x_k}(x_k)$ and $\varphi_{x_k}(v_x)$, where $v_x \in \mathbb{R}^n$. The goal is to determined the characteristic function of the propagated state at time step $k+1$, given by $x_{k+1} = \Phi x_k + \Gamma w_k$. It is assumed that x_k and the process noise $w_k \in \mathbb{R}^m$ are independent random vectors. The pdf and characteristic function of the latter are $f_W(w)$ and $\varphi_W(v_w)$, respectively, where $v_w \in \mathbb{R}^m$. $\Phi \in \mathbb{R}^{n\times n}$ and $\Gamma \in \mathbb{R}^{n\times m}$ are known matrices, with $|\Phi| \neq 0$. Due to the linearity of the state dynamics, the pdf of x_{k+1} is given as [14]

$$f_{x_{k+1}}(x_{k+1}) = \int_{-\infty}^{\infty} f_{x_k}\left(\Phi^{-1}x_{k+1} - \Phi^{-1}\Gamma w_k\right) f_W(w_k)\, dw_k. \quad \text{(C.1)}$$

Its characteristic function is given by

$$\varphi_{x_{k+1}}(v) = \Phi^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x_k}\left(\Phi^{-1}x_{k+1} - \Phi^{-1}\Gamma w_k\right) \times f_W(w_k)\, dw_k e^{jv^T x_{k+1}} dx_{k+1}. \quad \text{(C.2)}$$

Interchanging the order of integration, and using the substitution $x_{k+1} = \Phi x_k + \Gamma w_k \Rightarrow dx_{k+1} = |\Phi| dx_k$, the integral in (C.2) is solved as

$$\varphi_{x_{k+1}}(v) = \varphi_{x_k}(\Phi^T v)\, \varphi_W(\Gamma^T v). \quad \text{(C.3)}$$

C Appendix

Javier H. Fernández, Jason L. Speyer, and Moshe Idan
Linear Dynamic Systems with Additive Cauchy Noises Part 1: Stochastic Estimation for Two-State Systems
Submitted to IEEE Transactions on Automatic Control

Linear Dynamic Systems with Additive Cauchy Noises
Part 1: Stochastic Estimation for Two-State Systems

Javier H. Fernández *Student Member, IEEE*, Jason L. Speyer *Life Fellow, IEEE*, and Moshe Idan

***Abstract*—An efficient recursive state estimator is developed for two-state linear systems driven by Cauchy distributed process and measurement noises. For a general vector-state system, the estimator is based on recursively propagating the characteristic function of the conditional probability density function (cpdf), where the number of terms in the sum that expresses this characteristic function grows with each measurement update. Both the conditional mean and the conditional error variance are functions of the measurement history. The proposed two-state estimator reduces substantially the number of terms needed to express the characteristic function of the cpdf by taking advantage of relationships not yet developed in the general vector-state case. Further, by using a fixed sliding window of the most recent measurements, the improved efficiency of the proposed two state estimator allows an accurate approximation for real-time computation. In this way, the computational complexity of each measurement update eventually becomes constant, and an indefinite number of measurements can be processed. The performance of the Cauchy estimator was demonstrated numerically.**

I. Introduction

Dynamic processes involving uncertainty are frequently encountered in fields ranging from engineering and science to economics and finance. It is often assumed that the uncertainties are described by the Gaussian probability distribution, mainly because modern methods and algorithms are able to handle such systems very efficiently [1]. However, in many applications the underlying random processes have an impulsive character producing deviations of high amplitude and small duration much more often than the Gaussian assumption permits [2]. Examples of such processes include radar and sonar noise [3] and disturbances due to air turbulence [4].

Impulsive uncertainties were shown to be better described by heavy-tailed distributions, such as the symmetric alpha-stable ($S\alpha S$) distributions [5]. These distributions are described not by their probability density functions (pdfs), but by their characteristic functions (CFs). They are of the form $\varphi(v) = e^{-\sigma^\alpha |v|^\alpha + j\mu v}$, where σ is the scaling parameter, μ is the median, v is the spectral variable, and the characteristic exponent α determines the type of distribution: $\alpha = 2$ implies the Gaussian distribution, and $\alpha = 1$ implies the Cauchy distribution.

Estimation assuming Cauchy distributed noises has shown improved performance over Gaussian estimators when faced with impulsive noises. For estimating the direction of arrival of a signal to a sensor array in [6], maximum likelihood estimators designed assuming Cauchy distributed noises were shown to exhibit performance very close to the Cramèr-Rao Bound against $S\alpha S$ noises with characteristic exponents $1 \le \alpha \le 2$. Similar performance was observed in various applications, including processing data in a multi-user communication network [7] and radar glint [8]; in particular, the α parameter for the in-phase component of a time series of sea clutter in radar in [6] was calculated to be $\alpha \approx 1.7$. A framework based on these stable distribution models was developed in [9] and shown to have significant improvements in performance against heavy tailed noises.

The apparent robustness and adaptability of the Cauchy probability model motivated the derivation of a sequential estimator for linear scalar systems [10, 11], and subsequently for general vector-state systems driven by Cauchy noise [12, 13]. The estimator for general vector-state systems suffers from severe growth in numerical complexity, limiting its use to a small number of measurement updates and states. The aim of the current work is to develop an efficient two-state estimator to process measurements more quickly, and to arrest the growth in complexity in order to implement the Cauchy estimator for an arbitrary number of time steps.

The methodology in [12, 13] is based on finding the characteristic function of the conditional pdf of the state given the measurement history. This work follows that same procedure, and by exploiting certain relationships for the two-state structure, we can greatly reduce the complexity of the algorithm. The proposed algorithm was derived inductively by working out the first three measurement updates, and then deducing the general update process. In this paper, we present the first two measurement updates, followed by the general recursion. It has been checked against the results in [12, 13].

The CF is expressed as a sum of terms, each of which has two components: a coefficient function denoted by G and an exponential function with argument E. These functions are shown to have known structures that persist across measurement updates, and parameters that are contained in a set of fundamental arrays. The essence of this paper is in deriving this structure and populating the arrays, which allows for a drastic reduction in the complexity of the algorithm. We begin by presenting the problem in Section II, and performing the first measurement update for a general vector-state system in Section III. Then, in Section IV, we assume a two-state system and derive the proposed estimator structure for the first measurement update, followed by the first time propagation step in Section V. Section VI derives the proposed estimator structure for the second measurement update, and Section VII presents the general measurement update recursion algorithm. Section VIII discusses the finite horizon approximation of the full information estimator by using a fixed window of the most recent measurements. Finally, numerical examples are presented and discussed in Section IX, and concluding remarks are given in Section X.

II. Problem Formulation

Consider a discrete time, linear system described by

$$x(k+1) = \Phi x(k) + \Gamma w(k), \tag{1a}$$

$$z(k) = \mathrm{H} x(k) + v(k), \tag{1b}$$

where $x(k) \in \mathrm{R}^n$ is the state vector, $u(k)$ is a scalar deterministic input, $z(k)$ is a scalar measurement, and $w(k)$ and $v(k)$ are scalar independent Cauchy distributed process and measurement noise inputs. The noise inputs have medians at zero and scaling parameters

This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-09-1-0374, and by the United States - Israel Binational Science Foundation, Grant 2008040.

J. H. Fernández is with the department of Mechanical and Aerospace Engineering, University of California, Los Angeles.

Email: jhf@seas.ucla.edu.

J. L. Speyer is with the department of Mechanical and Aerospace Engineering, University of California, Los Angeles.

Email: speyer@seas.ucla.edu.

M. Idan is with the Faculty of Aerospace Engineering, Technion, Haifa, Israel. Email: moshe.idan@technion.ac.il.

of β and γ, respectively, so that their pdfs are given by

$$f_{W}(w(k)) = \frac{\beta/\pi}{w^2(k) + \beta^2}, \tag{2a}$$

$$f_{V}(v(k)) = \frac{\gamma/\pi}{v^2(k) + \gamma^2}. \tag{2b}$$

The characteristic functions of these pdfs are

$$\varphi_W(\sigma) = e^{-\beta|\sigma|}, \tag{3a}$$

$$\varphi_V(\sigma) = e^{-\gamma|\sigma|}, \tag{3b}$$

where σ is the scalar spectral variable.

The initial condition is assumed to be a product of scalar, independent Cauchy distributed random variables denoted by

$$f_{X_1}(x(1)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{(x_i(1) - \bar{x}_i(1))^2 + \alpha_i^2}. \tag{4a}$$

Its characteristic function is given by

$$\varphi_{X_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i |v_i| + j\bar{x}_i(1)v_i}$$

$$= \exp\left(-\sum_{i=1}^{n} \alpha_i |v_i| + j\bar{x}(1)^{\mathrm{T}} v\right), \tag{4b}$$

where v_i is an element of $v \in \mathbb{R}^n$. The algorithm's structure is greatly simplified by assuming that the median in (4b) is $\bar{x}(1) = 0$, i.e., the initial condition is centered at the origin. To preserve the generality of the initial condition (4b), the system in (1) can be decomposed into two systems, a stochastic variable initialized at the origin and driven by $w(k)$, and a deterministic system initialized at $\bar{x}(1)$, such that

$$x(k) = \bar{x}(k) + \tilde{x}(k), \tag{5a}$$

$$z(k) = \bar{z}(k) + \tilde{z}(k). \tag{5b}$$

The dynamics and measurement equations for $\tilde{x}$ and $\bar{x}$ are given by

$$\tilde{x}(k+1) = \Phi\tilde{x}(k) + w(k), \quad \tilde{z}(k) = \mathrm{H}\tilde{x}(k) + v(k), \tag{6a}$$

$$\bar{x}(k+1) = \Phi\bar{x}(k), \quad \bar{z}(k) = \mathrm{H}\bar{x}(k). \tag{6b}$$

Then, the proposed algorithm can be applied to the system in (6a), and the deterministic part in (6b) can be used to recover the state estimate for $x(k)$ from the estimate of $\tilde{x}(k)$ by using (5a). Therefore, the presentation of the algorithm will assume that $\bar{x}(1) = 0$ in (4b) without loss of generality.

Finally, the measurement history used in the estimation problem formulation is defined as

$$Z_k := \{z(1), \ldots, z(k)\}. \tag{7}$$

The objective is to derive the conditional mean estimator for this system. To do this, we find the CF of the pdf of the state at time k given the measurement history Z_k, denoted as $\varphi_{k|k}(v)$. The next section begins with the CF of the pdf conditioned on a single measurement.

III. First Measurement Update

Begin with the first measurement update at $k = 1$ by taking a noisy measurement of the Cauchy distributed initial state as

$$z(1) = \mathrm{H}x(1) + v(1). \tag{8}$$

Here, $x(1)$ and $v(1)$ are the Cauchy random variables. For a scalar system, i.e. $x(1) \in \mathbb{R}^1$, the conditional mean estimator has been derived and is presented in [10]. For vector-state systems, an approach based on determining the *characteristic function* (CF) of the conditional pdf (cpdf) is used. Initial results for estimation of scalar systems using the characteristic function of the conditional pdf were developed and presented in [11, 14]. An algorithm for a multivariate Cauchy estimator is presented in [12, 13] and summarized here.

The CF of the initial state conditioned on the first measurement is given by the vector integral

$$\varphi_{1|1}(v) = \int_{-\infty}^{\infty} f_{X_1|Z_1}(x(1)|z(1)) e^{jv^{\mathrm{T}}x(1)} dx(1). \tag{9}$$

The conditional pdf is computed from the joint distribution of $x(1)$ and $z(1)$ using Baye's Theorem [1] as

$$f_{X_1|Z_1}(x(1)|z(1)) = \frac{f_{X_1,Z_1}(x(1), z(1))}{f_{Z_1}(z(1))}$$

$$= \frac{f_{Z_1|X_1}(z(1)|x(1)) f_{X_1}(x(1))}{f_{Z_1}(z(1))}$$

$$= \frac{f_V(z(1) - \mathrm{H}x(1)) f_{X_1}(x(1))}{f_{Z_1}(z(1))}. \tag{10}$$

Then, (9) can be expressed as

$$\varphi_{1|1}(v) = \frac{1}{f_{Z_1}(z(1))} \int_{-\infty}^{\infty} f_V(z(1) - \mathrm{H}x(1)) \times f_{X_1}(x(1)) e^{jv^{\mathrm{T}}x(1)} dx(1) = \frac{\bar{\varphi}_{1|1}(v)}{f_{Z_1}(z(1))}, \tag{11}$$

where $\bar{\varphi}_{1|1}(v)$ is the characteristic function of the unnormalized cpdf (ucpdf). Note that, since $z(1)$ is known, $f_{Z_1}(z(1))$ is a constant; since $\varphi_{1|1}|_{v=0} = 1$, then $\bar{\varphi}_{1|1}|_{v=0} = f_{Z_1}(z(1))$.

Using the dual convolution property [15], $\bar{\varphi}_{1|1}(v)$ in (11) can be expressed as n convolution integrals in the v domain between the characteristic functions of $f_V(z(1) - \mathrm{H}x(1))$ and $f_{X_1}(x(1))$. The CF of $f_{X_1}(x(1))$ is given in (4b). The CF of $f_V(z(1) - \mathrm{H}x(1))$ is denoted $\hat{\varphi}_V(v)$ and is given by

$$\hat{\varphi}_V(v) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_V\Big(z(1) - \sum_{i=1}^{n} h_i x_i(1)\Big) \times e^{j\sum_{i=1}^{n} v_i x_i(1)} dx_1(1) \ldots dx_n(1). \tag{12}$$

In order to proceed, we need some assumptions about the measurement vector $\mathrm{H} \triangleq [h_1 \ldots h_n]$. The first is that at least one element of H is nonzero, i.e. there exists an i such that $h_i \neq 0$. This assumption is a prerequisite for observability of the state. The second assumption is that this nonzero element is h_n, which has no effect on generality [16].

To carry out the integration in (12), perform the change of variables: $\xi = z(1) - \sum_{i=1}^{n} h_i x_i$ to write

$$x_n(1) = \frac{1}{h_n}\left(z(1) - \xi - \sum_{i=1}^{n-1} h_i x_i(1)\right) \tag{13a}$$

$$dx_n(1) = \frac{d\xi}{|h_n|}. \tag{13b}$$

This allows us to manipulate (12) as

$$\hat{\varphi}_V(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_V(\xi)\times e^{j\left(\sum_{i=1}^{n-1} v_i x_i(1) + v_n \frac{z(1)-\xi-\sum_{i=1}^{n-1} h_i x_i(1)}{h_n}\right)} \frac{d\xi\, dx_1(1)\ldots dx_{n-1}(1)}{|h_n|}$$
$$= \frac{e^{j\frac{v_n}{h_n}z(1)}}{|h_n|}\int_{-\infty}^{\infty} f_V(\xi)e^{-j\frac{v_n}{h_n}\xi}d\xi \times \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{j\sum_{i=1}^{n-1}\left(v_i - \frac{h_i}{h_n}v_n\right)x_i(1)} dx_1(1)\ldots dx_{n-1}(1). \quad (14)$$

Here, the left parenthesis equals the CF of f_V if the spectral variable is $-\frac{v_n}{h_n}$. The right parenthesis is a product of Dirac delta functions, $\delta(\cdot)$, giving

$$\hat{\varphi}_V(v) = \frac{e^{j\frac{v_n}{h_n}z(1)}}{|h_n|}\varphi_V\left(-\frac{v_n}{h_n}\right)\prod_{i=1}^{n-1}(2\pi)\delta\left(v_i - \frac{h_i}{h_n}v_n\right). \quad (15)$$

Using φ_{X_1} of (4b) and $\hat{\varphi}_V$ of (15), we can express the first measurement update ucpdf's CF using the dual convolution property [15] as

$$\bar{\varphi}_{1|1}(v) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1}(x(1))f_V(z(1) - Hx(1))e^{jv^T x(1)}dx$$
$$= \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\varphi_{X_1}(v-\sigma)\hat{\varphi}_V(\sigma)d\sigma$$
$$= \frac{(2\pi)^{n-1}}{(2\pi)^n|h_n|}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\varphi_{X_1}(v-\sigma)e^{j\frac{\sigma_n}{h_n}z(1)}\varphi_V\left(-\frac{\sigma_n}{h_n}\right)\times\prod_{i=1}^{n-1}\delta\left(o_i - \frac{h_i}{h_n}\sigma_n\right)d\sigma_1\ldots d\sigma_{n-1}\,d\sigma_n. \quad (16)$$

Integrating over $\sigma_1\ldots\sigma_{n-1}$ is simple due to the delta functions in (16), and results in a single integral over the scalar σ_n:

$$\bar{\varphi}_{1|1}(v) = \frac{1}{2\pi|h_n|}\int_{-\infty}^{\infty}\varphi_{X_1}\left(v_1 - h_1\frac{\sigma_n}{h_n},\ldots,v_{n-1} - h_{n-1}\frac{\sigma_n}{h_n}, v_n - \sigma_n\right)\times e^{j\frac{\sigma_n}{h_n}z(1)}\varphi_V\left(-\frac{\sigma_n}{h_n}\right)d\sigma_n$$
$$= \frac{1}{2\pi|h_n|}\int_{-\infty}^{\infty}\varphi_{X_1}\left(v - H^T\frac{\sigma_n}{h_n}\right)e^{j\frac{\sigma_n}{h_n}z(1)}\varphi_V\left(-\frac{\sigma_n}{h_n}\right)d\sigma_n. \quad (17)$$

Finally, we have a convolution integral involving the CF of the pdf of the state, and the CF of the measurement. This result indicates that the ucpdf's CF for a system of arbitrary (finite) order conditioned on a scalar measurement can be determined from this single, scalar convolution integral. A simple change of scalar variables $\sigma = \sigma_n/h_n$ and $d\sigma_n = d\sigma\,|h_n|$ gives the final form

$$\bar{\varphi}_{1|1}(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi_{X_1}(v - H^T\sigma)\,\varphi_V(-\sigma)\,e^{jz(1)\sigma}d\sigma. \quad (18)$$

Rewrite (18) using (3b) and (4b) as

$$\bar{\varphi}_{1|1}(v) = \int_{-\infty}^{\infty}\left(\prod_{i=1}^{n+1}\right)\frac{1}{2\pi}\exp\left(-\sum_{i=1}^{n+1}\rho_i|\mu_i - \sigma| + jz(1)\sigma\right)d\sigma, \quad (19a)$$

$$\rho_i = \alpha_i - E_i H^T \text{ for } i \in \{1,\ldots,n\} \qquad \rho_{n+1} = \gamma, \quad (19b)$$

$$\mu_i = \frac{E_i v}{E_i H^T} \text{ for } i \in \{1,\ldots,n\} \qquad \mu_{n+1} = 0, \quad (19c)$$

where E_i is the i^{th} row of the n-dimensional identity matrix. Note that the μ_i are scalars linear in v, i.e. inner products of v with given vectors. The solution to (19) is given in [12, 13] as

$$\bar{\varphi}_{1|1}(v) = \frac{1}{2\pi}\sum_{i=1}^{n+1}G_{1|1}^{i}(v)\cdot e^{E_i\Gamma_{1|1}(v)}$$
$$= \frac{1}{2\pi}\sum_{i=1}^{n+1}\exp\left(-\sum_{\substack{l=1\\ l\neq i}}^{n+1}\rho_l|\mu_l - \mu_i| + jz(1)\mu_i\right)\times\left(jz(1) + \rho_i + \sum_{\substack{l=1\\ l\neq i}}^{n+1}\rho_l\,\mathrm{sgn}(\mu_l - \mu_i)\right)^{-1}$$
$$-\left(jz(1) - \rho_i + \sum_{\substack{l=1\\ l\neq i}}^{n+1}\rho_l\,\mathrm{sgn}(\mu_l - \mu_i)\right)^{-1}. \quad (20)$$

Thus, the measurement update process produces a sum of exponential terms with coefficients in the brackets. For the two state system, as will be shown, the bracket term can be reduced to a simple form that is a polynomial of sign functions. Next, we present a structure for expressing (20) for the two-state system that can be extended to subsequent measurement updates.

IV. Two-State Estimator Structure for the First Measurement Update

In the previous section we considered a system with a state vector of general order n, and an algorithm for the general vector-state system is presented in [12, 13]. However, that algorithm suffers from a very aggressive growth in computational complexity with each new measurement update. For a second order system there are certain patterns and algebraic relationships that allow for significant reductions in numerical complexity and allow the estimator to run effectively over a large number of measurement updates. There are two main aspects to this simplification: a way of expressing and indexing vectors that multiply (as inner products) the spectral vector v in the exponential argument in (20); and a set of algebraic relations that can be used to simplify the coefficients of the exponential functions. Both of these aspects are addressed here for the first measurement update in (20) presented above. In Section VI the second update is presented, indicating by induction the general measurement update and time propagation recursions given in Section VII.

A. Exponential Argument

Consider the arguments of the absolute value terms in (20). The μ scalars are defined in (19c) as scaled inner products of v with vectors we call *fundamental directions*. For the first measurement update, these fundamental directions are the rows of the $n\times n$ identity matrix. For the two-state system, the set of fundamental directions from the

initial condition is $B_{1|0} = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. In the subscript of $B_{1|0}$, the first element denotes the time step, and the second element denotes the number of measurements that have been processed.

Inner products are linear operations, and hence a difference of inner products with a given vector is also an inner product. This new inner product introduces a new fundamental direction. This new direction is the same for any real, 2×2 $B_{1|0}$. Using a superscript on $B_{1|0}$ to denote its i^{th} row, apply this notation to the definitions in (19c) to get

$$\mu - \mu_m = \frac{B_{1|0}v}{B_{1|0}\mathrm{H}^{\mathrm{T}}} - \frac{B^m_{1|0}v}{B^m_{1|0}\mathrm{H}^{\mathrm{T}}} = \frac{\mathrm{H}B^{m\mathrm{T}}_{1|0}B_{1|0}v - \mathrm{H}B^{\mathrm{T}}_{1|0}B^m_{1|0}v}{\mathrm{H}B^{m\mathrm{T}}_{1|0}B_{1|0}\mathrm{H}^{\mathrm{T}}} = \frac{\mathrm{H}\left(B^{m\mathrm{T}}_{1|0}B_{1|0} - B^{\mathrm{T}}_{1|0}B^m_{1|0}\right)v}{(B^m_{1|0}\mathrm{H}^{\mathrm{T}})(B_{1|0}\mathrm{H}^{\mathrm{T}})}. \tag{21}$$

The key here is to recognize that the term in parenthesis in the numerator of (21) is a matrix minus its own transpose, which implies that it is antisymmetric, i.e., that $B^{m\mathrm{T}}_{1|0}B_{0} - B^{\mathrm{T}}_{1|0}B^m_{1|0} = \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}$ for some $c \in \mathbb{R}$. This constant c can be computed and pulled out of the matrix, which allows us to express any two-dimensional antisymmetric matrix as $c\Lambda$ where $\Lambda = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and c can be verified to be

$$c = -B_{1|0}\Lambda B^{m\mathrm{T}}_{1|0}. \tag{22}$$

Hence, we can write (21) as

$$\mu - \mu_m = \left(\frac{-B_{1|0}\Lambda B^{m\mathrm{T}}_{1|0}}{(B^m_{1|0}\mathrm{H}^{\mathrm{T}})(B_{1|0}\mathrm{H}^{\mathrm{T}})}\right) \cdot \mathrm{H}\Lambda v. \tag{23}$$

This produces the new fundamental direction is $\mathrm{H}\Lambda$, scaled by the term in parenthesis. For the relationship in (23) to hold, $\mu \neq \mu_m$ and neither μ nor μ_m can equal zero. In (19c) we also defined an extra constant $\mu_3 = 0$. This implies that the old fundamental directions are retained in the measurement updated cpdf's CF for terms in the exponent of (20) that involve μ_3. Therefore, the set of fundamental directions for the first measurement update, denoted by $B_{1|1}$, is given by

$$B_{1|1} = \begin{bmatrix} E_1 \\ E_2 \\ \mathrm{H}\Lambda \end{bmatrix}. \tag{24}$$

In (20), we use (23) to express $\mu_2 - \mu_1$ as

$$\mu_2 - \mu_1 = \frac{E_2 v}{E_2\mathrm{H}^{\mathrm{T}}} - \frac{E_1 v}{E_1\mathrm{H}^{\mathrm{T}}} = \frac{-(E_2\Lambda E_1^{\mathrm{T}})\mathrm{H}\Lambda v}{(E_1\mathrm{H}^{\mathrm{T}})(E_2\mathrm{H}^{\mathrm{T}})} = \frac{\mathrm{H}\Lambda v}{\gamma}. \tag{25}$$

Using (25) and the definitions for ρ from (19b) in (20) yields

$$\begin{aligned}
\varphi_{1|1}(v) = {} & G^1_{1|1}(v)\exp\left(-\alpha_2 E_2\mathrm{H}^{\mathrm{T}} - \frac{\mathrm{H}\Lambda v}{(E_1\mathrm{H}^{\mathrm{T}})(E_2\mathrm{H}^{\mathrm{T}})}\right. \\
& \left. -\gamma\frac{E_1 v}{E_1\mathrm{H}^{\mathrm{T}}} + j z(1)\frac{E_1 v}{E_1\mathrm{H}^{\mathrm{T}}}\right) \\
& + G^2_{1|1}(v)\exp\left(-\alpha_1 E_1\mathrm{H}^{\mathrm{T}} - \frac{\mathrm{H}\Lambda v}{(E_1\mathrm{H}^{\mathrm{T}})(E_2\mathrm{H}^{\mathrm{T}})}\right. \\
& \left. -\gamma\frac{E_2 v}{E_2\mathrm{H}^{\mathrm{T}}} + j z(1)\frac{E_2 v}{E_2\mathrm{H}^{\mathrm{T}}}\right) \\
& + G^3_{1|1}(v)\exp\left(-\alpha_1 E_1\mathrm{H}^{\mathrm{T}}\frac{E_1 v}{E_1\mathrm{H}^{\mathrm{T}}}\right. \\
& \left. -\alpha_2 E_2\mathrm{H}^{\mathrm{T}}\frac{E_2 v}{E_2\mathrm{H}^{\mathrm{T}}}\right) \\
= {} & G_{1|1}(v) \\
& \times \exp\left(-\frac{\gamma}{|E_1\mathrm{H}^{\mathrm{T}}|}|E_1 v| - \frac{\alpha_2}{|E_1\mathrm{H}^{\mathrm{T}}|}|\mathrm{H}\Lambda v| + j\frac{z(1)}{E_1\mathrm{H}^{\mathrm{T}}}E_1 v\right) \\
& + G^2_{1|1}(v) \\
& \times \exp\left(-\frac{\gamma}{|E_2\mathrm{H}^{\mathrm{T}}|}|E_2 v| - \frac{\alpha_1}{|E_2\mathrm{H}^{\mathrm{T}}|}|\mathrm{H}\Lambda v| + j\frac{z(1)}{E_2\mathrm{H}^{\mathrm{T}}}E_2 v\right) \\
& + G^3_{1|1}(v)\exp\left(-\alpha_1|E_1 v| - \alpha_2|E_2 v|\right).
\end{aligned} \tag{26}$$

Notice that each of the three terms involves only two of the three fundamental directions.

The efficiency of the proposed two-state estimator is achieved by both keeping track of which directions are used in each term, as well as the scalar coefficients that multiply the absolute value functions and the scalar coefficients in the imaginary part of the argument of the exponential. The most important of these is an array of integers where the elements of each row correspond to the rows of $B_{1|1}$ that appear in the exponential argument. This array is denoted by $M(1|1)$.

The exponential argument of the term corresponding to $i = 3$ in (26) is exactly the same as the initial condition. In fact, the only difference between the initial condition and this term is the coefficient $G^3_{1|1}$. This is due to the cancellations that occur in general in (26), so that one term produced from the convolution will always have the same exponential argument with a different coefficient function. This term is referred to as the *old term*. It will be shown later that the old exponential arguments persist across measurement updates, and therefore it is useful to order the terms with the old fundamental directions first, so that the last term in (26) moves to the top. Consequently, $M(1|1)$ is constructed as

$$M(1|1) = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix}. \tag{27}$$

The other two terms, corresponding to $i = 1$ and $i = 2$ in (26), are called the *intermediate new terms*; we say *intermediate* because in all subsequent measurement updates, as will be shown, they will combine with other terms with the same exponential argument, and *new* because they involve the new fundamental directions just generated during the measurement update.

Following this new ordering, we define two additional arrays $P(1|1)$ and $Z(1|1)$ whose elements correspond to the coefficients of the absolute value functions in the exponents and the coefficients in the imaginary part, respectively; hence, the i^{th} rows of $P(1|1)$ and $Z(1|1)$ are related to the i^{th} term in $\varphi_{1|1}$. For the measurement

$(\boldsymbol{E}_1\mathrm{H}^{\mathrm{T}})(\boldsymbol{E}_2\mathrm{H}^{\mathrm{T}})$ updated ordering of terms used to define $\boldsymbol{M}(1\,|\,1)$ in (27), these arrays

are given by

$$P(1|1) = \begin{bmatrix} \frac{\alpha_1}{|E_1 H^T|} & \frac{\alpha_2}{|E_1 H^T|} \\ \frac{\gamma}{|E_2 H^T|} & \frac{\alpha_1}{|E_2 H^T|} \end{bmatrix}, \quad Z(1|1) = \begin{bmatrix} \frac{z_1}{E_1 H^T} & 0 \\ 0 & \frac{z_1}{E_2 H^T} \end{bmatrix}. \tag{28}$$

They have the same dimension as $M(1|1)$, so that an element of $P(1|1)$ or $Z(1|1)$ goes with the fundamental direction indexed by the corresponding element of $M(1|1)$.

Finally, define a vector array of integers $L_{1|1}$ with as many elements as rows of $M(1|1)$. Each element of $L_{1|1}$ indicates the number of fundamental directions in that term, i.e., the width of the corresponding row of $M(1|1)$. For the first measurement update, $L_{1|1}$ is given by

$$L_{1|1} = \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}^T. \tag{29}$$

This final array is unnecessary in the first measurement update, but will become essential later when different terms involve different numbers of fundamental directions. Finally, define the number of terms as $N_{1|1}$. Clearly, $N_{1|1} = 3$, the same as the number of rows of $M(1|1)$, $P(1|1)$, $Z(1|1)$, and $L_{1|1}$.

Using (24) (27) (28) (29) and the definition for $N_{1|1} = 3$, the exponential argument for the first measurement update ucpdf's CF given in (20) can be expressed as

$$E_{1|1}(v) = -\sum_{i=1}^{L_{1|1}^i} P_i \left(B_{1|1}^{M_i} v\right)^2 + j \sum_{i=1}^{L_{1|1}^i} z_i B_{1|1}^{M_i} v, \quad i \in \{1, \ldots, N_{1|1}\}. \tag{30}$$

B. Polynomial Coefficient

Using (25), (19b), and (19c), the bracketed terms $G_{1|1}^i$ in (20) are of the same form, and each involves two sign functions, denoted s_1 and s_2, with the same arguments as the absolute values in the corresponding exponential parts. Functions of this form can be reduced to a four parameter polynomial of these sign functions as $a_i + b_i s_1 s_2 + j c_i s_1 + j d_i s_2$. These relationships are given in Appendix A; for $k = 1$ in particular, Result 3 in Appendix A can be used to obtain the four parameter polynomial by assuming that $a_i^m = 1$ and $b_i^m = c_i^m = d_i^m = 0$. The sign functions involve the same fundamental directions as the exponential argument, so denote a final array $G(1|1)$ with the same number of rows as $M(1|1)$ and width four, and let each row contain the parameters for the polynomial, as

$$G(1|1) = \begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{bmatrix}. \tag{31}$$

Note that the same reordering used to form $M(1|1)$ must be used here, so that the top row contains the parameters for the old term's coefficient.

Using the parameters in (31), along with (24) and (27), the coefficient function $G_{1|1}^i$ for the first update ucpdf's CF given in (20) can be expressed as a polynomial

$$G_{1|1}^i(v) = a_i + b_i \operatorname{sgn}\left(B_{1|1}^{M_i^1} v\right) \operatorname{sgn}\left(B_{1|1}^{M_i^2} v\right) + j c_i \operatorname{sgn}\left(B_{1|1}^{M_i^1} v\right) + j d_i \operatorname{sgn}\left(B_{1|1}^{M_i^2} v\right). \tag{32}$$

Remark 1: In the notation, for P, Z, and M the subscripts denote the row (i.e. which term in the sum) and the superscripts denote the that is indexed by the integer M_i. It is assumed that P, Z, and M inherit their time index from the associated fundamental directions from B.

V. First Time Propagation

In this section the characteristic function of the cpdf is propagated using the dynamics given in (1a). Assume that at an arbitrary time step k the state's pdf is given by $f_{x_k}(x(k))$, its CF is given by $\varphi_{k|k}(v)$, and that the process noise CF is given as in (3a). The pdf of $x(k+1)$ is determined using the linear transformation of the joint distribution $f_{x_k,w}(x_k, w_k)$. The details of this derivation can be found in [12, 13], and the formula for the characteristic function of the propagated cpdf is given by

$$\varphi_{k+1|k}(v) = \varphi_{k|k}(\Phi^T v)\varphi_w(\Gamma^T v). \tag{33}$$

Note that the derivation of (33) in [12, 13] requires that Φ be invertible.

Applying the time propagation equation (33) to the characteristic function of the first measurement update given in (20) yields

$$\varphi_{2|1}(v) = \varphi_{1|1}(\Phi^T v)\varphi_w(\Gamma^T v) = \frac{\bar{\varphi}_{1|1}(\Phi^T v)\varphi_w(\Gamma^T v)}{f_Z(z(1))} = \frac{\bar{\varphi}_{2|1}(v)}{f_{Z_1}(z(1))}. \tag{34}$$

Then, the ucpdf's characteristic function for the propagated state is given by

$$\bar{\varphi}_{2|1}(v) = \bar{\varphi}_{1|1}(\Phi^T v)\varphi_w(\Gamma^T v) = \bar{\varphi}_{1|1}(\Phi^T v) \cdot e^{-\beta|\Gamma^T v|} = \sum_{i=1}^{N_{1|1}} G^i \exp\left(E_{1|1}^i(\Phi^T v) - \beta|\Gamma^T v|\right)$$

$$= \sum_{i=1}^{N_{1|1}} G^i \exp\left(\sum_{i=1}^{L_{1|1}^i} P_i \left(B_{1|1}^{M_i} \Phi^T v\right)^2 - \beta|\Gamma^T v|\right) \cdot \exp\left(j \sum_{i=1}^{L_{1|1}^i} z_i B_{1|1}^{M_i} \Phi^T v\right). \tag{35}$$

The changes in $\bar{\varphi}_{2|1}(v)$ from $\bar{\varphi}_{1|1}(v)$ are a new element in the sum of the absolute value terms, and a linear transformation on v. The time propagation step adds no new terms to the sum, hence $N_{2|1} = N_{1|1}$. Moreover, since the process noise has a zero median, the time propagation has no effect on the complex part of the exponential argument, so that $Z(2|1) = Z(1|1)$. Finally, there is no effect on the parameters of the coefficient functions G, so that $G(2|1) = G(1|1)$. The time propagated ucpdf's characteristic function can be restated in the same form as $\varphi_{1|1}$, i.e.,

$$\bar{\varphi}_{2|1}(v) = \sum_{i=1}^{N_{2|1}} G_{2|1}^i(v) \cdot e^{E_{2|1}^i(v)}, \tag{36a}$$

where

$$G_{2|1}^i(v) = a_i + b_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} v\right) \operatorname{sgn}\left(B_{2|1}^{M_i^2} v\right) + j c_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} v\right) + j d_i \operatorname{sgn}\left(B_{2|1}^{M_i^2} v\right), \tag{36b}$$

column (i.e., where in the exponential argument it appears). Only B, L, and N retain the time dependence in their subscripts. The superscript for B denotes the row, i.e., the fundamental direction,

$$E^{i}_{2|1}(v) = -\sum_{=1}^{L_{2|1}} P_i \, B^{M_i}_{2|1} v + j \sum_{=1}^{} Z_i B^{M_i}_{2|1} v, \qquad (36c)$$

$$Z(2|1) = Z(1|1),$$

$$P(2|1) = \begin{bmatrix} \beta & \frac{\alpha_1}{\beta} & \frac{\alpha_2}{\beta} \\ \gamma & \frac{\alpha_2}{|E_1 H^{\mathrm{T}}|} & \beta \\ \beta & \frac{\gamma}{|E_2 H^{\mathrm{T}}|} & \frac{\alpha_1}{|E_2 H^{\mathrm{T}}|} \end{bmatrix} P(1|1) \begin{bmatrix} \beta & \frac{\alpha_2}{|E_1 H^{\mathrm{T}}|} & \beta \\ \gamma & \frac{\alpha_1}{|E_2 H^{\mathrm{T}}|} & \beta \end{bmatrix},$$

$$B_{2|1} = \begin{bmatrix} B_{1|1} \Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} E_1 \Phi^{\mathrm{T}} \\ E_2 \Phi^{\mathrm{T}} \\ H \Lambda \Phi^{\mathrm{T}} \end{bmatrix}, \tag{36d}$$

$$M(2|1) = \begin{bmatrix} 4 & 4 & 4 \\ 4 & & \end{bmatrix} M(1|1) \begin{bmatrix} 4 & 4 \\ 4 & \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix},$$

$$L_{2|1} = L_{1|1} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

The process noise introduces a third absolute value term, with a new fundamental direction, into all of the CF's exponential arguments. The next section deals with performing the next measurement update to this propagated structure.

VI. Second Measurement Update

The second measurement update process involves finding the characteristic function of the unnormalized conditional pdf, i.e., $\varphi_{2|2}$.

The formula used for this second measurement update is the same as the first update, given in (18), and is applied to (36):

$$\begin{aligned} \varphi_{2|2}(v) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_{2|1}(v - H^{\mathrm{T}}\sigma) \varphi_V(-\sigma) e^{j\sigma z(2)} d\sigma \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\varphi}_{2|1}(v - H^{\mathrm{T}}\sigma) e^{-\gamma|\sigma|} e^{j\sigma z(2)} d\sigma. \end{aligned} \tag{37}$$

Since $\bar{\varphi}_{2|1}$ is a sum of $N_{2|1} = 3$ terms, the measurement update process is to solve the convolution integral $N_{2|1}$ times, once for each term. Substituting (36a) into (37), we get

$$\begin{aligned} \varphi_{2|2}(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{i=1}^{N_{2|1}} & G^i_{2|1}(v - H^{\mathrm{T}}\sigma) \\ & \times \exp\left[E^i_{2|1}(v - H^{\mathrm{T}}\sigma) - \gamma|-\sigma| + j\sigma z(2) \right] d\sigma. \end{aligned} \tag{38}$$

Interchanging the integration and summation operations produces $N_{2|1}$ convolution integrals. As with the first measurement update, each convolution will produce an old term as well as new terms, called *intermediate terms*. Since many of these terms have the same exponential arguments, they can be combined to reduce the total number of terms in the CF's sum.

Therefore, we begin with an arbitrary i^{th} convolution in (38). Using (36c) we define

$$\begin{aligned} I_i = & \int_{-\infty}^{\infty} G^i_{2|1}(v - H^{\mathrm{T}}\sigma) \\ & \times \exp\left[E^i_{2|1}(v - H^{\mathrm{T}}\sigma) - \gamma|\sigma| + j\sigma z(2) \right] \\ = & \int_{-\infty}^{\infty} G^i_{2|1}(v - H^{\mathrm{T}}\sigma) \\ & \times \exp\left[-\sum_{i=1}^{L^i_{2|1}} P_i \, B^{M_i}_{2|1}(v - H^{\mathrm{T}}\sigma) - \gamma|\sigma| \right] \\ & \times \exp\left[j \sum_{i=1}^{2} Z^i_i B^{M^i_i}_{2|1}(v - H^{\mathrm{T}}\sigma) + j\sigma z(2) \right] d\sigma. \end{aligned} \tag{39}$$

In order to solve (39), we need to rewrite it in a manner similar to (19) using ρ and μ substitutions. Begin as in the first measurement update by defining constants μ obtained by expressing

$$B^{M^i_{2|1}}_{2|1}(v - H^{\mathrm{T}}\sigma) = B^{M^i_{2|1}}_{2|1}{}^{\mathrm{T}}(\mu - \sigma), \tag{40a}$$

and thus

$$\mu = \frac{B_{M^i_{2|1}} v}{B^{M^i_{2|1}}_{2|1} H^{\mathrm{T}}}. \tag{40b}$$

Next, we rewrite the argument of the exponential of (39) in terms of these μ s and σ as

$$\begin{aligned} & -\sum_{i=1}^{L^i_{2|1}} P_i \, B^{M_i}_{2|1} H^{\mathrm{T}} \cdot |\mu - \sigma| - \gamma|-\sigma| \\ & \quad + j \left(Z^1_i B^{M_{i1}}_{2|1} + Z^2_i B^{M_{i2}}_{2|1} \right) v \\ & \quad + j \left(z(2) - Z^1_i B^{M_{i1}}_{2|1} H^{\mathrm{T}} - Z^2_i B^{M_{i2}}_{2|1} H^{\mathrm{T}} \right) \sigma \\ & = j \left(Z^1_i B^{M^1_i}_{2|1} + Z^2_i B^{M^2_i}_{2|1} \right) v \\ & \quad - \sum_{i=1}^{L^i_{2|1}+1=4} \rho_i |\mu_i - \sigma| + j\theta^i_2 \sigma, \end{aligned} \tag{41a}$$

where the complex part multiplying v does not depend on σ and comes out of the convolution. The parameter definitions are $\mu_{L^i_{2|1}+1} = \mu_4 = 0$,

$$\theta^i_2 = z(2) - Z^1_i B^{M^1_i}_{2|1} H^{\mathrm{T}} - Z^2_i B^{M^2_i}_{2|1} H^{\mathrm{T}}, \tag{41b}$$

and the $L^i_{2|1} + 1 = 3 + 1 = 4$ constants called ρ are given by

$$\rho_i = \begin{cases} \gamma & i \in \{1, \ldots, L^i_{2|1} = 3\}, \\ B^{M_i}_{2|1} \left[\begin{smallmatrix} \\ \\ \end{smallmatrix} \right]_{i \in \{1, \ldots, N_{2|1} = 3\}} & i = L^i_{2|1} + 1 = 4. \end{cases} \tag{41c}$$

This is the same procedure used in the first measurement update except that, due to the new absolute value term in the exponential argument introduced in the time propagation step, there are now four ρ constants instead of three. Moreover, there are three fundamental directions in the exponential argument, instead of just two.

Let's turn our attention to the coefficients of the exponents. The coefficient functions for the three terms in $\varphi_{2|1}$ are all of the form

$$\begin{aligned} G^i_{2|1}(v) = a_i + b_i \operatorname{sgn}\left(B^i_{2|1} v \right) \operatorname{sgn}\left(B^{M_{i2}}_{2|1} v \right) + \\ j c_i \operatorname{sgn}\left(B^{M^1_i}_{2|1} v \right) + j d_i \operatorname{sgn}\left(B^{M^2_i}_{2|1} v \right). \end{aligned} \tag{42}$$

Hence the same manipulation used in (40b) to form the μs can be used here to get

$$G^i_{2|1}(v - \mathrm{H}^{\mathrm{T}}\sigma) = a_i + b_i \operatorname{sgn}\left(B^{M^1_i}_{2|1}\mathrm{H}^{\mathrm{T}} \cdot B^{M^2_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right) \times \left(\frac{1}{2}\left(\operatorname{sgn}(\mu_1 - \sigma) + \operatorname{sgn}(\mu_2 - \sigma)\right) - 1\right) + jc_i \operatorname{sgn}\left(B^{M^1_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}(\mu_1 - \sigma) + jd_i \operatorname{sgn}\left(B^{M^2_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}(\mu_2 - \sigma), \quad (43)$$

where the identity $\operatorname{sgn}(\mu_1 - \sigma)\operatorname{sgn}(\mu_2 - \sigma) = \frac{1}{2}\left(\operatorname{sgn}(\mu_1 - \sigma) + \operatorname{sgn}(\mu_2 - \sigma)\right) - 1$ is used.

In order to use the same integration method as in the first measurement update, it is necessary for the coefficient of the exponential term to be constant within each subdomain of integration $(\mu_\ell, \mu_{\ell+1})$. Since these coefficients are polynomials of sums of sign functions, they are clearly constant in these regions. Now define the following constants,

$$\bar{a}_i = a_i, \qquad \bar{b}_i = b_i \operatorname{sgn}\left(B^{M^1_i}_{2|1}\mathrm{H}^{\mathrm{T}} \cdot B^{M^2_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right),$$
$$\bar{\rho}_1 = c_i \operatorname{sgn}\left(B^{M^1_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right), \qquad \bar{\rho}_2 = d_i \operatorname{sgn}\left(B^{M^2_i}_{2|1}\mathrm{H}^{\mathrm{T}}\right),$$
$$\rho_1 = 1, \qquad \rho_2 = 1. \quad (44)$$

Our integration method also requires that the coefficient involve sums over the same sign functions as the exponential argument, which has two more absolute value functions than the coefficient has sign functions. To do this, introduce two more sign functions into the sum in (43) by multiplying them by constants $\rho_3 = \rho_4 = 0$ and $\bar{\rho}_3 = \bar{\rho}_4 = 0$. Using these, the coefficient in the convolution can be written with sums over the same μ as in (41a):

$$\begin{aligned} G^i_{2|1}(v - \mathrm{H}^{\mathrm{T}}\sigma) &= \bar{a}_i + \bar{b}_i \left(\frac{1}{2}\left(\sum_{\ell=1}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right) - 1\right) + j \sum_{\ell=1}^{4} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma) \\ &= G^i_{2|1}\left(\sum_{\ell=1}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \sigma), \sum_{\ell=1}^{4} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right) \\ &:= G^i_{2|1}\left(\sum_{\ell=1}^{4} \rho \operatorname{sgn}(\mu_\ell - \sigma)\right). \quad (45) \end{aligned}$$

Since the only difference between the summations in (45) are the scalar ρ and $\bar{\rho}$ constants, we can simplify the notation and use a shorthand ρ in (45) to represent all the summations. This notation will become more useful in subsequent measurement updates as the coefficient polynomials become more complex.

Using these substitutions, each of the $i \in \{1, \ldots, N_{2|1}\}$ convolution integrals in (38) can be written as

$$I_i = \int_{-\infty}^{\infty} G^i_{2|1}\left(\sum_{\ell=1}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \sigma)\right) \times \exp\left(-\sum_{\ell=1}^{4} \rho_\ell |\mu_\ell - \sigma| + j\theta^i_2 \sigma\right) d\sigma \quad (46)$$

ment update from [12, 13] is

$$I_i = L^{i}_{2|1+1} \exp\left(-\sum_{m=1}^{L^i_{2|1}+1} \rho_m |\mu_m - \mu_m| + j\theta^i_2 \mu_m\right) \times \frac{G^i_{2|1}\left(\sum_{m=1}^{L^i_{2|1}+1} \rho_m + \sum_{\ell=1, \ell \neq m}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right), \quad j\theta^i_2 + \sum_{m=1}^{L^i_{2|1}+1} \rho_m + \sum_{\ell=1, \ell \neq m}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right) - G^i_{2|1}\left(-\sum_{m=1}^{L^i_{2|1}+1} \rho_m + \sum_{\ell=1, \ell \neq m}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right), \quad j\theta^i_2 - \sum_{m=1}^{L^i_{2|1}+1} \rho_m + \sum_{\ell=1, \ell \neq m}^{L^i_{2|1}+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right)}{}. \quad (47)$$

The convolution produces $L^i_{2|1} + 1 = 4$ terms, indexed here by m, for each of the i terms in the sum in $\bar{\varphi}_{2|1}$. The first three terms are called the *intermediate new terms* because some of the terms from different convolutions have the same exponential arguments and can be combined by summing their coefficient functions together, as will be shown. The last term produced by the convolution that corresponds to $m = 4$ is the *old term* because, since $\mu_4 = 0$, it has the same exponential argument as the i^{th} term in $\bar{\varphi}_{2|1}$ that was convolved. Simplifications of these terms and their coefficient functions are addressed next.

A. Second Measurement Update - New Intermediate Terms

The terms corresponding to $m \in \{1, 2, 3\}$ are called the *intermediate new terms*. These always involve only two fundamental directions: one is from $B_{2|1}$ and the other is the new direction HA as in the first measurement update. The formation of the new fundamental direction is what recovers the structure involving two fundamental directions, as will be shown here.

Until the intermediate new terms are combined into the final structure for the second measurement update, it is useful to denote the exponential argument in terms of both indices i and m. Hence, for the intermediate terms, denote the coefficient and exponential arguments as $G^{i,m}$ and $E^{i,m}$, respectively.

2|2 2|2

The solution to the i^{th} convolution integral for the second measure-

For the i^{th} term, use the definitions in (40b), (41b), and (41c) to rewrite the exponential argument of (47) as

$$E^{i,m}_{2|2} = -\gamma \frac{B^{M^m_i}_{2|1} v}{B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}}} - \sum_{\ell=1}^{L^i_{2|1}+1} P^{i}_{\ell} B^{M_i}_{2|1} \mathrm{H}^{\mathrm{T}} \frac{-B^{M_i}_{2|1} \Lambda B^{M^m_i}_{2|1}}{B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}} \cdot B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}}} |\mathrm{H}\Lambda v| + j \frac{z(2)}{B^{M^m_i}_{2|1} \mathrm{H}_{\mathrm{T}} B^{M^m_i}_{2|1} v} + j \sum_{\ell=1}^{2} Z^{i}_{\ell} B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}} \frac{-B^{M_i}_{2|1} \Lambda B^{M^m_i}_{2|1}}{B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}} \cdot B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}}} \mathrm{H}\Lambda v - \gamma \frac{B^{M^m_i}_{2|1}}{B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}}} \sum_{\ell=1}^{L^i_{2|1}+1} P^{i}_{\ell} B^{M_i}_{2|1} \Lambda B^{M^m_i}_{2|1} |\mathrm{H}\Lambda v|$$

$$= \frac{z(2) B^{M^m_i}_{2|1} v - j \sum_{\ell=1}^{2} Z^{i}_{\ell} B^{M_i}_{2|1} \Lambda B^{M^m_i}_{2|1} \mathrm{H}\Lambda v}{B^{M^m_i}_{2|1} \mathrm{H}^{\mathrm{T}}} + j \frac{}{}. \quad (48)$$

This final form for the new intermediate term's exponential argument involves only two fundamental directions: one that is a row of $B_{2|1}$ that corresponds to the M^m_i integers for $m \in \{1, \ldots, L^i_{2|1} = 3\}$, and the new vector $\mathrm{H}\Lambda$. Moreover, the denominators of both terms are equal. Therefore, we define the set of updated fundamental directions as $B_{2|2}$ and construct it by appending $\mathrm{H}\Lambda$ to the bottom of $B_{2|1}$ as

$$B_{2|2} = \begin{bmatrix} B_{2|1} \\ \mathrm{H}\Lambda \end{bmatrix} = \begin{bmatrix} E_1\Phi^{\mathrm{T}} \\ E_2\Phi^{\mathrm{T}} \\ \mathrm{H}\Lambda\Phi^{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \\ \mathrm{H}\Lambda \end{bmatrix}. \quad (49)$$

B. Second Measurement Update - Recursive Structure of the Arrays

Denote an array of integers $M(2|2)$ where the elements of a given row index the rows of $B_{2|2}$ that appear in the corresponding term. The rows of the old terms will be unchanged, and a set of new rows will be appended to the bottom, each of width two. Since each convolution produces terms involving one of the old term's directions, the new rows of $M(2|2)$ will contain all possible combinations of rows of $B_{2|1}$ with the new row in $B_{2|2}$ as

$$M(2|1) = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$$

$$M(2|2) = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \\ 4 & 5 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & 5 \\ 3 & 5 \\ 4 & 5 \\ 4 & 5 \\ 4 & 5 \end{bmatrix}. \quad (50)$$

Each of the three convolutions in the second measurement update produced four terms, the old term and three new intermediate ones, for a total of nine new terms. However, from (50) it is clear that we only have six distinct new terms with different pairs of fundamental directions. This is because three pairs of terms, from different convolutions, have the same exponential arguments and can be combined into one term by summing their polynomial coefficients. same fundamental directions, the coefficients for the absolute value functions are different, as will be shown next.

Now we construct an array of coefficients for the absolute value functions in the exponents. From (48), a pattern emerges for constructing these coefficients. They are stored in the array $P(2|2)$ that is given by

$$P(2|2) = \begin{bmatrix} \frac{\gamma}{|E_1\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|} & \frac{P(2|1)}{\alpha_2 |\det\Phi| + \beta E_1\Phi^{\mathrm{T}}\Lambda\Gamma} \\ & \frac{|E_1\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|}{\alpha_1 |\det\Phi| + \beta E_2\Phi^{\mathrm{T}}\Lambda\Gamma} \\ \frac{\gamma}{|E_2\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|} & \frac{|E_2\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|}{\gamma \det\Phi + \beta \mathrm{H}\Lambda\Phi^{\mathrm{T}}\Lambda\Gamma} \\ \frac{\gamma}{|\mathrm{H}\Lambda\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|} & \frac{|\mathrm{H}\Lambda\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}|}{\alpha_1 E_1\Phi^{\mathrm{T}}\Lambda\Gamma + \alpha_2 E_2\Phi^{\mathrm{T}}\Lambda\Gamma} \\ \frac{\gamma}{|\mathrm{H}\Gamma|} & \frac{|\mathrm{H}\Gamma|}{\gamma E_1\Phi^{\mathrm{T}}\Lambda\Gamma + \frac{\alpha_2}{|\mathrm{H}^{\mathrm{T}}|} \mathrm{H}\Lambda\Phi^{\mathrm{T}}\Lambda\Gamma} \\ \frac{\gamma}{|\mathrm{H}\Gamma|} & \frac{|\mathrm{H}\Gamma|}{E_2\Phi^{\mathrm{T}}\Lambda\Gamma + \frac{\alpha_1}{|\mathrm{H}^{\mathrm{T}}|} \mathrm{H}\Lambda\Phi^{\mathrm{T}}\Lambda\Gamma} \\ & |\mathrm{H}\Gamma| \end{bmatrix}. \quad (51)$$

Next, denote the new array of coefficients for the imaginary part of the exponential argument as $Z(2|2)$. Since the time propagation step has no effect on the imaginary part of the exponential argument, this array always has width two, involving only the original two directions, and it has the same number of rows as $M(2|2)$. Based on the same manipulations used to obtain the elements of $P(2|2)$, $Z(2|2)$ is given by

$$Z(2|2) = \begin{bmatrix} Z(2|1) & \\ z(2) & 0 \\ E_1\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}} & \\ z(2) & 0 \\ E_2\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}} & \\ z(2) & -z(1)\det\Phi \\ \mathrm{H}\Lambda\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}} & \\ \frac{z(2)}{\mathrm{H}\Gamma} & 0 \\ \frac{z(2)}{\mathrm{H}\Gamma} & -\frac{z(1) \cdot E_1\Phi^{\mathrm{T}}\Lambda\Gamma}{E_1\mathrm{H}^{\mathrm{T}}\mathrm{H}\Gamma} \\ \frac{z(2)}{\mathrm{H}\Gamma} & -\frac{z(1)}{E_2\mathrm{H}^{\mathrm{T}} \cdot E_2\Phi^{\mathrm{T}}\Lambda\Gamma} \end{bmatrix}. \quad (52)$$

The usefulness of $L_{2|2}$ is more apparent in this measurement update, since older terms involve more fundamental directions. It is formed by simply appending an array of 2s of length six to $L_{2|1}$ as

$$L_{2|2} = \begin{bmatrix} L^{\mathrm{T}}_{2|1} & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 3 & 3 & 3 & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}}. \quad (53)$$

C. Second Measurement Update - New Coefficients G

Consider now the new coefficient functions for the new intermedi-

These three combined pairs produce the first three new terms, and thus they introduce the first three new rows of $M(2|2)$. The last three rows of $M(2|2)$ are due to the time propagation step and involves the Γ^{-T} direction. Although these last three terms have the ate terms produced by the i^{th} term of $\bar{\varphi}$. For $m = 1$ and $m = 2$, the numerators of (47) are not equal and hence cannot come out of the bracket term. They are of a form compatible with Result 3 in Appendix A. Denote the coefficient of the m^{th} intermediate term as

$G_{2|1}^{i,m}$. Then, the numerators for $m = 1$ are given by

$$\begin{aligned} G_{2|1}^{i,1}(\cdot) &= \pm\rho_1 + \sum_{\ell=2}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_1) \\ &= \bar{a}_i + \bar{b}_i \left(\frac{1}{2} \left(\pm\rho_1 + \sum_{\ell=2}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_1) \right)^2 - 1 \right) \\ &\quad + j \left(\pm\bar{\rho}_1 + \sum_{\ell=2}^{4} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_1) \right) \\ &= \bar{a}_i + \bar{b}_i \left(\frac{1}{2} \left(\pm\rho_1 + \rho_2 \operatorname{sgn}(\mu_2 - \mu_1) \right)^2 - 1 \right) \\ &\quad + j \left(\pm\bar{\rho}_1 + \bar{\rho}_2 \operatorname{sgn}(\mu_2 - \mu_1) \right) \\ &= a_i \pm b_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} \mathrm{H}^{\mathrm{T}} \cdot B_{2|1}^{M_i^2} \mathrm{H}^{\mathrm{T}} \right) \operatorname{sgn}(\mathrm{HA}v) \\ &\quad \pm j c_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} \mathrm{H}^{\mathrm{T}} \right) + j d_i \operatorname{sgn}\left(B_{2|1}^{M_i^2} \mathrm{H}^{\mathrm{T}} \right) \operatorname{sgn}(\mathrm{HA}v). \end{aligned} \tag{54}$$

The same manipulations can be done for the numerators for $m = 2$ by interchanging μ_1 and c_i with μ_2 and d_i, respectively. Recalling that $\rho_3 = 0$, the new term obtained for $m = 3$ corresponds to the coefficient

$$\begin{aligned} G_{2|1}^{i,3}(0) &= \sum_{\ell=1}^{4} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_3) = \\ &a_i + b_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} \mathrm{H}^{\mathrm{T}} \cdot B_{2|1}^{M_i^2} \mathrm{H}^{\mathrm{T}} \right) \operatorname{sgn}(\mu_1 - \mu_3) \operatorname{sgn}(\mu_2 - \mu_3) \\ &\quad + j c_i \operatorname{sgn}\left(B_{2|1}^{M_i^1} \mathrm{H}^{\mathrm{T}} \right) \operatorname{sgn}(\mu_1 - \mu_3) + \\ &\quad j d_i \operatorname{sgn}\left(B_{2|1}^{M_i^2} \mathrm{H}^{\mathrm{T}} \right) \operatorname{sgn}(\mu_2 - \mu_3) \\ &= a_i + b_i \operatorname{sgn}\left(-B_{2|1}^{M_i^1} \Lambda B_{2|1}^{M_i^3 \mathrm{T}} \right) \operatorname{sgn}\left(-B_{2|1}^{M_i^1} \Lambda B_{2|1}^{M_i^3 \mathrm{T}} \right) \\ &\quad + j c_i \operatorname{sgn}\left(-B_{2|1}^{M_i^1} \Lambda B_{2|1}^{M_i^3 \mathrm{T}} \right) \\ &\quad + j d_i \operatorname{sgn}\left(-B_{2|1}^{M_i^2} \Lambda B_{2|1}^{M_i^3 \mathrm{T}} \right) \operatorname{sgn}(\mathrm{HA}v). \end{aligned} \tag{55}$$

Substituting these into the bracket term in (47), and noting that the real parts of the denominators are similar to the real part of the exponential argument, yields a form that is compatible with the numerator forms given in Result 3 to produce coefficient functions for the intermediate terms denoted $G_{2|2}^{i,m}(v)$ and given by

$$\begin{aligned} G_{2|2}^{i,m}(v) = a_{i,m} + b_{i,m} \operatorname{sgn}\left(B_{2|2}^{M_{i,m}^1} v \right) \operatorname{sgn}\left(B_{2|2}^{M_{i,m}^2} v \right) + \\ j c_{i,m} \operatorname{sgn}\left(B_{2|2}^{M_{i,m}^1} v \right) + j d_{i,m} \operatorname{sgn}\left(B_{2|2}^{M_{i,m}^2} v \right). \end{aligned} \tag{56}$$

It is necessary to use Result 3 because the denominators in the bracket term in (47) form two sign functions, in the same way that the new exponential arguments (48) involve only two absolute value terms. Therefore, polynomial coefficients for all of the new intermediate terms can be expressed in the simple four-parameter form of (56). Note that the two fundamental directions involved here are the same as in the argument of the exponential. Hence, combining two terms with the same exponential arguments involves summing their polynomial coefficients $G_{2|2}^{i,m}$, which is simply summing the corresponding parameters.

For the second update, three pairs of the nine new intermediate terms can be combined, leaving a total of six new terms in addition to the three old terms. The set of combined parameters for the $G(1|1)$, producing the new array $G(2|2)$, which has four columns and the same number of rows as $M(2|2)$. Then, the coefficients for the new terms (i.e.,those corresponding to $i \in \{4, \ldots, 9\}$) are given by

$$\begin{aligned} G_{2|2}^{i}(v) = a_i + b_i \operatorname{sgn}\left(B_{2|2}^{M_i^1} v \right) \operatorname{sgn}\left(B_{2|2}^{M_i^2} v \right) \\ + j c_i \operatorname{sgn}\left(B_{2|2}^{M_i^1} v \right) + j d_i \operatorname{sgn}\left(B_{2|2}^{M_i^2} v \right), \end{aligned} \tag{57}$$

where a_i, b_i, c_i, and d_i are the elements of the i^{th} row of $G(2|2)$.

D. Second Measurement Update - Old Terms

For $i = L_{k|k-1}^{i} + 1$ the constants $\{\rho_\ell, \bar{\rho}_\ell\} := \rho_1 = 0$, which implies that both numerators in the bracket term in (47) are equal and hence can be pulled out of the brackets as a common factor. Therefore, the measurement updated old term is the same as the previous old term, except its coefficient is multiplied by the new bracket term,

$$\begin{aligned} G_{2|2}^{i}(v) = \left(a_i + b_i \operatorname{sgn}\left(B_{2|2}^{M_i^1} v \right) \operatorname{sgn}\left(B_{2|2}^{M_i^2} v \right) \right. \\ \left. + j c_i \operatorname{sgn}\left(B_{2|2}^{M_i^1} v \right) + j d_i \operatorname{sgn}\left(B_{2|2}^{M_i^2} v \right) \right) \\ \times \frac{1}{2\pi} \left| j\theta_2^i + \gamma + \sum_{\ell=1}^{i_{2|2}=3} P_\ell^i \left(B_{2|2}^{M_i} \mathrm{H}^{\mathrm{T}} \operatorname{sgn}\left(B_{2|2}^{M_i} v \right) \right)_{-1} \right| \\ - \left| j\theta_2^i + \gamma + \sum_{\ell=1}^{i_{2|2}=3} P_\ell^i \left(B_{2|2}^{M_i} \mathrm{H}^{\mathrm{T}} \operatorname{sgn}\left(B_{2|2}^{M_i} v \right) \right)_{-1} \right|. \end{aligned} \tag{58}$$

Since the new bracket term in (58) has three sign functions, it cannot be manipulated using the results in Appendix A.

This implies that terms that have been created in previous measurement updates retain their exponential arguments during the current measurement update, and their coefficients are multiplied by new bracket terms as in (58), with one additional sign function than the coefficient had in the previous update. The difference between the current time step and the time step when the term was originally created is called the *age* of the term, and is given by $L_{k|k} - 2$. Although this complicated structure grows, acquiring a new bracket term every measurement update, the structure of each bracket term overlaps significantly with the others, and thus can be expressed efficiently.

E. Second Measurement Update - The Characteristic Function

Using (49) (50) (51) (52) (53), and denoting $N_{2|2} = 9$ for the number of terms, the CF of the ucpdf for the second measurement update is given by

$$\varphi_{2|2}(v) = \sum_{i=1}^{N_{2|2}} G_{2|2}^{i}(v) \cdot \exp\left(E_{2|2}^{i}(v) \right) \tag{59a}$$

where

$$E_{2|2}^{i}(v) = -\sum_{\ell=1}^{L_{2|2}^{i}} P_\ell^i \left(B_{2|2}^{M_i} v \right)_{-2} + j \sum_{\ell=1}^{i_{2|2}} Z_\ell B^{M_i} v \tag{59b}$$

and the coefficients are given by (57) and (58) for new and old terms, respectively. Note that the first three terms in $\varphi_{2|2}$ correspond to the old terms and have the same exponential parts as the three terms in $\varphi_{2|1}$. The subsequent six terms are new and involve only two fundamental directions. Although not used here, it will be useful to denote the number of new terms in $\varphi_{2|2}$ as $N''_{2|2} = 6$. This structure second measurement update are appended to the bottom of $G(2|1) =$

repeats itself across subsequent measurement updates, and the next section will present the two-state estimator's recursion for the general k^{th} measurement update.

VII. Estimator Recursion for the k^{th} Update and Propagation

This section presents the two-state Cauchy estimator algorithm for a general measurement update k. This algorithm as derived by induction based on a study of the first three measurement updates, of which two are shown here. It can be verified to be a special case of the approach developed in [12, 13]. We first address the update at time step k, and then follow with the time propagation step. Finally, we show how to determine the minimum variance estimate and its error variance.

A. Measurement Update

We begin this section assuming we have a time propagated ucpdf's CF $\varphi_{k|k-1}$ and are given the k^{th} measurement $z(k)$. Consider the following generalization for the polynomial coefficients $G^i_{k|k-1}$ produced for the old terms in the second measurement update. The form in (58) suggests that, in general, the time propagated coefficients will have the form

$$G^i_{k|k-1}(v) = \left[a_i + b_i \operatorname{sgn}\left(B^1_{k|k-1}v\right) \operatorname{sgn}\left(B^2_{k|k-1}v\right) + jc_i \operatorname{sgn}\left(B^1_{k|k-1}v\right) + jd_i \operatorname{sgn}\left(B^2_{k|k-1}v\right)\right] \times \prod_{r=1}^{L^i_{k|k-1}-3} \frac{1}{2\pi} \frac{1}{j\theta^i_{k-r} - \gamma + S^i_{k-r}(B_{k|k-1}v)} - \frac{1}{j\theta^i_{k-r} - \gamma + S^i_{k-r}(B_{k|k-1}v)}, \quad (60a)$$

where

$$S^i_{k-r}(B_{k|k-1}v) = \sum_{i=1}^{L^i_{k-r|k-r}} P_i \left(B^{M_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}\left(B^{M_i}_{k|k-1}v\right), \quad (60b)$$

$$\theta^i_{k-r} = z(k-r) - Z_i B^{M^1_{k-r|k-r}}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} - Z_i B^{M^2_{k-r|k-r}}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} \quad (60c)$$

and $k-r$ is the time-step where the term involving S^i_{k-r} was created. The arrays M,L,P,B, and Z for the general update will be constructed later but correspond to those of the first two updates already shown.

Using the measurement update formula (18) with the corresponding change of indices, $\varphi_{k|k}$ is given by

$$\bar{\varphi}_{k|k}(v) = \sum_{i=1}^{N_{k|k-1}} \frac{1}{2\pi} \int_{-\infty}^{\infty} G^i_{k|k-1}(v - \mathrm{H}^{\mathrm{T}}\sigma) \times \exp\left[-\sum_{i=1}^{L^i_{k|k-1}} P_i \left(B^{M_i}_{k|k-1}\right)(v - \mathrm{H}^{\mathrm{T}}\sigma) - \gamma|-\sigma| + j\theta^i_k\sigma\right] d\sigma. \quad (61)$$

It is necessary to divide the domain of integration into regions in which the polynomial coefficient is a constant and the exponential argument is continuous. For this, the summations that appear in (60) must involve the same fundamental directions as the exponential argument. Since the top bracket term in (60a) is the four parameter polynomial formed when the term was created, we can define $\bar{\rho}$ and $\bar{\rho}_m$ constants as in (44), where $\bar{\rho} = \rho = 0$ for $f, = 3, \ldots, L^i_{k|k-1} + 1$.

Similarly, the summations in (60b), from the rest of the bracket terms in (60a), can be written using constants ${}^{(r)}\rho$ defined as

$${}^{(r)}\rho = \sum_{i=0}^{} P_i \left(B^{M}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}\left(B^{M}_{k|k-1}\mathrm{H}^{\mathrm{T}}\right) \quad (62a)$$

where

$$f, = L^i f, = 1, \ldots, L^i_{k|k-1} - r \quad i \quad k|k-1 - r + 1, \ldots, L_{k|k-1} + 1 \quad (62b)$$

Using the shorthand ρ as in the second measurement update (45), construct μ , θ^i_k, ρ , $\bar{\rho}$, and $\bar{\rho}$ constants as in (40b), (41b), (41c), and (44), respectively, in order to write the coefficients in the i^{th} integral in (61) as

$$\begin{aligned} G^i_{k|k-1}(v - \mathrm{H}^{\mathrm{T}}\sigma) &= \\ &G^i_{k|k-1}\Big[\sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \sigma), \sum_{=1}^{L^i_{k|k-1}+1} \bar{\rho} \operatorname{sgn}(\mu - \sigma), \\ &\sum_{=1}^{L^i_{k|k-1}+1} {}^{(L^i_{k|k-1}-2)}\rho \operatorname{sgn}(\mu - \sigma), \\ &\ldots, \sum_{=1}^{L^i_{k|k-1}+1} {}^{(1)}\rho \operatorname{sgn}(\mu - \sigma)\Big] \\ &:= G^i_{k|k-1}\Big[\sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \sigma)\Big]. \end{aligned} \quad (63)$$

Then, the entire integral can be written compactly as

$$I_i = \int_{-\infty}^{\infty} G^i_{k|k-1}\Big[\sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \sigma)\Big] \times \exp\Big[-\sum_{=1}^{L^i_{k|k-1}+1} \rho |\mu - \sigma| + j\theta^i_k\sigma\Big] d\sigma. \quad (64)$$

The solution to this integral, given in [12, 13], is

$$I_i = \sum_{m=1}^{L^i_{k|k-1}+1} \exp\Big[-\sum_{=1}^{L^i_{k|k-1}+1} \rho |\mu - \mu_m| + j\theta_k\mu_m\Big] \times \frac{G^i_{k|k-1}\Big[+\rho_m + \sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \mu_m)\Big]}{j\theta^i_k + \rho_m + \sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \mu_m)} - \frac{G^i_{k|k-1}\Big[-\rho_m + \sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \mu_m)\Big]}{j\theta^i_k - \rho_m + \sum_{=1}^{L^i_{k|k-1}+1} \rho \operatorname{sgn}(\mu - \mu_m)}, \quad (65a)$$

where

$$G^i_{k|k-1}(\pm\bar{\rho}_m + \sum_{\ell=1}^{L^{i\,k|k-1}_{k|k-1}} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)) =$$

$$\left(a_i + b_i \left(\pm\bar{\rho}_m + \sum_{\ell=1}^{L^{i\,k|k-1}_{k|k-1}} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right)^{-1}\right)^{-2}$$

$$\times \prod_{\ell=1,\ell\neq m}^{L^{i\,k|k-1}_{k|k-1}} \left(\pm\bar{\rho}_m + \sum_{\ell=1}^{L^{i\,k|k-1}_{k|k-1}} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\right)$$

$$\times \prod_{p=1}^{L^{i\,k|k-1}_{k-r}-3} \frac{1}{2\pi j\theta^i_{k-r} + \gamma + S^i_{k-r}(\pm^{(r)}\bar{\rho}_m, \mu_\ell - \mu_m)} - \frac{1}{j\theta^i_{k-r} - \gamma + S^i_{k-r}(\pm^{(r)}\bar{\rho}_m, \mu_\ell - \mu_m)}, \quad (65b)$$

and for $r \in \{1, \ldots, L^{i\,k|k-1}_{k-r} - 3\}$,

$$S^i_{k-r}(\pm^{(r)}\bar{\rho}_m, \mu_\ell - \mu_m) = \pm P^{m\,M_m}_{i\,k-r|k-r} B^{M_m}_{k-r|k-r} \mathrm{H}^{\mathrm{T}} + \sum_{\ell=1,\ell\neq m}^{L^{i\,k|k-1}_{k-r}} \left(\pm P^{m\,M_m}_{i\,k-r|k-r} B^{M_m}_{k-r|k-r} \mathrm{H}^{\mathrm{T}} + B^{M_\ell}_{k|k-1}\mathrm{H}^{\mathrm{T}}\right) \operatorname{sgn}(\mu_\ell - \mu_m)$$

$$= \pm P^{m\,M_m}_{i\,k-r|k-r} B^{M_m}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} + \sum_{\ell=1}^{L^{i\,k|k-1}_{k-r}} P^{m\,M_m}_{i\,k-r|k-r} B^{M_m}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} \operatorname{sgn}(-B^{M_\ell}_{k|k-1}\mathrm{A}B^{M_\ell}_{k|k-1} \cdot B^{M_\ell}_{k|k-1}\mathrm{H}^{\mathrm{T}}) \times \operatorname{sgn}(\mathrm{HA}v). \quad (65c)$$

B. Measurement Update - Recovering the CF Structure

The old terms will be discussed later. For the new terms, i.e., $m = 1, \ldots, L^i_{k|k-1}$, the numerators of the measurement updated coefficients (65b) can be reduced to a form compatible with Result 3. The four parameter bracket terms in (65b) can be rewritten as (56) from the second measurement update. Manipulate the bracket term outside the product in (65b) in the same manner as (54) and (55). Then, the entire numerator form (65b) can be collapsed, using Results 1 and 2, into the form given in (56). Finally, the new four-parameter polynomial coefficient involving two fundamental directions (HA and a row of $B_{k|k-1}$) can be computed as in the second measurement update using Result 3. Those parameters will be combined with other terms with the same exponential arguments, the set of new parameters will be appended to the bottom of $G(k|k-1)$ to form a new array denoted $\bar{G}(k|k)$. Denote the number of the new terms in $\varphi_{k|k}$ as $N^n_{k|k}$ so that $N_{k|k} = N_{k|k-1} + N^n_{k|k}$.

Since one of the two fundamental directions for every new term will be HA, it is appended to $B_{k|k-1}$ to obtain

$$B_{k|k} = \begin{bmatrix} M(k|k-1) & 2k+1 \\ M^1_{(r_k+1)} & \\ M^1 & \end{bmatrix}$$

$$M(k|k) = \begin{bmatrix} M^{(r_k+2)}_{k} & 2k+1 \\ \vdots & \\ M^1_{N^n_{k|k}} & 2k+1 \\ \vdots & \vdots \\ 2k & 2k+1 \end{bmatrix}, \quad (67)$$

where $r_k = N_{k-1|k-1} - N^n_{k-1|k-1}$, noting that $N_{k|k-1} = N_{k-1|k-1}$ and $N^n_{k|k-1} = N_{k-1|k-1}$. The left column of this new block has three parts. The first $N^n_{k-1|k-1}$ elements are the left elements of all the new rows of the previous measurement update, i.e. the new $N^n_{k-1|k-1}$ rows of $M(k-1|k-1)$. The next element corresponds to the new fundamental direction from the previous measurement update, i.e. the last row of $B_{k-1|k-1}$. The remaining $N_{k-1|k-1}$ rows are all the same, and involve the new time propagation direction Γ^{T}.

The growth in the number of terms, which is given as a linear dynamic system of integers, is based on the pattern for the recursion of $M(k|k)$ given above. The number of terms in the sum in $\varphi_{k|k}$, given by $N_{k|k}$, is determined from the previous number of terms $N_{k-1|k-1}$ and the previous number of new terms $N^n_{k-1|k-1}$ by the following linear relationship:

$$\begin{bmatrix} N_{k|k} \\ N^n_{k|k} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} N_{k-1|k-1} \\ N^n_{k-1|k-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (68)$$

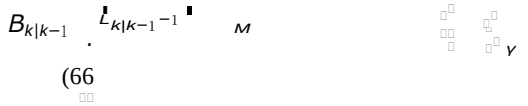
The recursions for the fundamental arrays $P(k|k)$ and $Z(k|k)$ were derived by induction from a study of the first three measurement updates. As in the first two measurement updates, the denominators of the elements of the i^{th} row of $Z(k|k)$ are equal. Therefore, it is useful to denote the diagonal matrix $D(k|k)$, where

$$D(k|k) = \operatorname{Diag}\left(\frac{1}{B^{M^1_{(r_k+1)}}_{k|k}\mathrm{H}^{\mathrm{T}}}, \ldots, \frac{1}{B^{M^1_{(r_k+N^n_{k|k-1})}}_{k|k}\mathrm{H}^{\mathrm{T}}}, \frac{1}{\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{H}^{\mathrm{T}}}, \frac{1}{\mathrm{H}\Gamma}, \ldots, \frac{1}{\mathrm{H}\Gamma}\right). \quad (69)$$

Similarly, the denominators of the elements of the i^{th} row of $P(k|k)$ are also equal. Denote the diagonal matrix $D(k|k)$, whose elements equal the absolute values of the corresponding elements in $D(k|k)$.

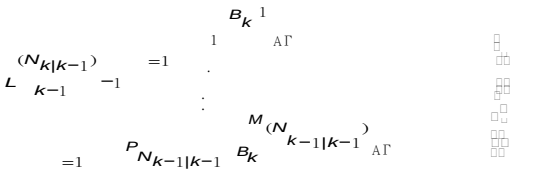
The new terms in $P(k|k)$ and $Z(k|k)$ are formed in the same manner as in the first and second measurement update, reducing sums of $\mu - \mu_m$ terms to constants times $\mathrm{H}\Lambda$ and reducing the polynomial coefficients to the four parameter structure in (23) using the Results in Appendix A, producing the new rows for $G(k|k)$. The recursion for $P(k|k)$ is

$$P(k|k) = \bar{D}(k|k) \times \begin{bmatrix} P(k|k-1) & \\ P^2_{(r_k+1)} & B^{M^1_{(r_k+1)}}_{k-1|k-1}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi + \beta\, B^{M^1_{(r_k+1)}}_{k|k}\mathrm{A}\Gamma \\ P^2_{(r_k+2)} & B^{M^1_{(r_k+2)}}_{(r_k+2)\,k-1|k-1}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi + \beta\, B^{M^1_{(r_k+2)}}_{k|k}\mathrm{A}\Gamma \\ \vdots & \vdots \\ P^2_{r_k+N^n_{k|k-1}} & B^{M^1_{(r_k+N^n_{k|k-1})}}_{k-1|k-1}\mathrm{H}^{\mathrm{T}} \cdot \det\Phi + \beta\, B^{M^1_{(r_k+N^n_{k-1|k-1})}}_{k|k}\mathrm{A}\Gamma \\ \gamma|\det\Phi| + \beta\,\mathrm{HA}\Phi^{\mathrm{T}}\mathrm{A}\Gamma & \end{bmatrix}.$$



(66

)



The array $M(k|k)$ is constructed by appending $N_{k|k}^{n}$ new rows to $M(k|k-1)$. The recursion is given by

(70)

Similarly, the recursion for $Z(k|k)$ is

$$Z(k|k) = \begin{bmatrix} D(k|k)\times & Z(k|k-1) & \\ z(k) & 0 & \\ z(k) & 0 & \\ z(k) & -Z^2_{(r_k+3)}\left(B^{(r_k+3)}_{k-1|k-1}\right)\mathrm{H}^{\mathrm{T}}\cdot\det\Phi & \\ \vdots & \vdots & \\ z(k) & -Z^2_{(r_k+N^n_{k-1})}\left(B^{(r_k+N^n_{k-1})}_{k-1|k-1}\right)\mathrm{H}^{\mathrm{T}}\cdot\det\Phi & \\ z(k) & -z(k-1)\cdot\det\Phi & \\ z(k) & -\sum_{=1}^{2} Z_{(N_{k-1|k-1})}\left(B^{M^1_{k|k}}_{k|k}\right)\Lambda\Gamma & \\ \vdots & \vdots & \\ z(k) & -\sum_{=1}^{2} Z_{(N_{k-1|k-1})}\left(B^{M(N_{k-1|k-1})}_{k|k}\right)\Lambda\Gamma & \end{bmatrix}. \quad (71)$$

Since all the new rows in $M(k|k)$ have width two, $L_{k|k}$ is measurement updated by appending $N^n_{k|k}$ elements of the integer 2 to $L_{k|k-1}$ as

$$L^{\mathrm{T}}_{k|k} = \begin{bmatrix} L_{k|k-1}, & 2, & \ldots, & 2 \end{bmatrix}^{\mathrm{T}}. \quad (72)$$

This produces all of the parameters necessary to express the ucpdf's CF for the k^{th} measurement update as

$$\bar{\varphi}_{k|k}(v) = \sum_{i=1}^{N_{k|k}} G^i_{k|k}(v) \times \exp\left[-L^i_{k|k}\sum_{=1} P_i\left(B^{M_i}_{k|k}v\right) + j\sum_{=1}^{2} Z_i B^{M_i}_{k|k}v\right], \quad (73)$$

where coefficients are given by

$$G_{k|k}(v) = a_i + b_i \operatorname{sgn}\left(B^{M^1_i}_{k|k}v\right)\operatorname{sgn}\left(B^{M^2_i}_{k|k}v\right) + jc_i \operatorname{sgn}\left(B^{M^1_i}_{k|k}v\right) + jd_i \operatorname{sgn}\left(B^{M^2_i}_{k|k}v\right). \times \prod_{r=1}^{L^i_{k|k}-2}\left[\frac{1}{2\pi}\frac{1}{j\theta^i_{k-r}+\gamma+S^i_{k-r}(B_{k|k}v)} - \frac{1}{j\theta^i_{k-r}-\gamma+S^i_{k-r}(B_{k|k}v)}\right], \quad (74a)$$

where

$$S^i_{k-r}(B_{k|k}v) = \sum_{=1}^{L^i_{k-r|k-r}} P_i\left(B^{M_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}\right)\operatorname{sgn}\left(B^{M_i}_{k|k}v\right), \quad (74b)$$

$$\theta^i_{k-r} = z(k-r) - Z_i B^{M^1_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}} - Z_i B^{M^2_i}_{k-r|k-r}\mathrm{H}^{\mathrm{T}}, \quad (74c)$$

and $k-r$ is the time-step where the term involving S^i_{k-r} was created. Hence, the age of the i^{th} term after the k^{th} update is $L^i_{k|k}-2$, and new terms have age zero.

C. Time Propagation

The general time propagation uses the same formula as in the first time propagation,

$$\bar{\varphi}_{k+1|k}(v) = \bar{\varphi}_{k|k}(\Phi^{\mathrm{T}}v)\cdot e^{-\beta\,\Gamma^{\mathrm{T}}v}. \quad (75)$$

$$\bar{\varphi}_{k+1|k}(v) = \sum_{i=1}^{N_{k+1|k}} G_{k+1|k}(v) \times \exp\left[-L^i_{k+1|k}\sum_{=1} P_i\left(B^{M_i}_{k+1|k}v\right) + j\sum_{=1}^{2} Z_i B^{M_i}_{k+1|k}v\right], \quad (76)$$

where

$$P(k+1|k) = \begin{bmatrix} P(k|k) & \beta & & \\ & \ddots & & \\ & & \beta & \\ & & & 2(k+1) \end{bmatrix}, \quad Z(k+1|k) = Z(k|k),$$

$$M(k+1|k) = \begin{bmatrix} M(k|k) & 2(k+1) \\ \vdots & 2(k+1) \end{bmatrix},$$

$$L_{k+1|k} = L_{k|k} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad B^{k+1|k} = \begin{bmatrix} B_{k|k}\cdot\Phi_{\mathrm{T}} \\ \Gamma^{\mathrm{T}} \end{bmatrix}, \quad (77)$$

and both $N_{k+1|k} = N_{k|k}$ and $N^n_{k+1|k} = N^n_{k|k}$ because no new terms are created during the time propagation.

D. Evaluating the Conditional Mean and Estimation Error Variance

It is shown in [12, 13] that $\bar{\varphi}_{k|k}(v)$ is twice continuously differentiable. The mean can be found from the unnormalized characteristic function [1] by taking its partial derivative and then taking the limit as v goes to the origin.

$$\hat{x}(k) = E[x(k)|Z_k] = jf_{Z_k}(Z_k)^{-1}\left(\frac{\partial\bar{\varphi}_{k|k}(v)}{\partial v}\right)^{\mathrm{T}}_{v\to 0}, \quad (78a)$$

where

$$f_{Z_k}(Z_k) = \bar{\varphi}_{k|k}(v)_{v\to 0}. \quad (78b)$$

The second moment can be found by taking the same limit of the second partial derivative of the characteristic function, as

$$E\left[x(k)x(k)^{\mathrm{T}}|Z_k\right] = f_{Z_k}(Z_k)^{-1}\frac{\partial^2\bar{\varphi}_{x_k|Z_k}(v)}{\partial v v^{\mathrm{T}}}_{v\to 0}, \quad (78c)$$

and the error variance is given by

$$\Xi(k) = E\left[x(k)x(k)^{\mathrm{T}}|Z_k\right] - \hat{x}(k)\hat{x}(k)^{\mathrm{T}}. \quad (78d)$$

The limits above must be taken along valid directions, due to the structure of $\varphi_{k|k}(v)$ in (76). Since this structure is a special case of the structure in [12, 13], the details of these operations can be found there.

VIII. Finite Horizon Approximation

In order to arrest the growth in computational complexity, we approximate the full information CF with one using a fixed sliding window of the most recent measurements, where number of measurements in this horizon is denoted N_Z. Hence, the first N_Z measurement updates in the estimation are performed normally. Then, for every measurement update $k > N_Z$, we initialize a new finite horizon (FH) estimator and perform N_Z measurement updates over the fixed window $\{z(k-N_Z+1), \ldots, z(k)\}$. This new initial condition for the FH estimator is of the form

$$\varphi^k_{w_{1|0}}(v) = \exp\left[-\alpha_1\left(B_{w_{1|0}}v\right)_w - \alpha_2\left(B_{w_{1|0}}v\right)_w\right]$$

$$\begin{matrix} \mathbf{r} & & +j\bar{x} \\ & {}_{1} & v_1 \\ \cos\phi & \sin\phi & \mathbf{l} \\ & j\bar{x}_2 & \\ & v_2 & , \end{matrix}$$

The characteristic function of the ucpdf for the once propagated conditional density can be written as

$$B_{w1|0} = \begin{matrix} & \\ -\sin\phi & \cos\phi \end{matrix}, \tag{79}$$

where $B_{W1|0}$ is a rotation matrix. Denote the windowed first measurement update mean and variance by $\hat{x}_k^W(1)$ and $\Xi_k^W(1)$. The

1 2 1 2

of $\varphi_{k-N_Z+1|k-N_Z+1}$ and of the updated FH first update $\varphi_{W1|1}^k$. This process involves solving five nonlinear equations with the five unknowns stated above, which is carried out using standard numerical tools. It is necessary to make use of the decomposition in (5a) in order to apply the proposed algorithm to the initial condition in (79).

This local first measurement updated CF, then, has the same mean and variance as the original CF we are approximating. The remaining $N_Z - 1$ measurement updates are performed over the measurements in the window, ultimately producing the N_Z-measurement updated CF $\varphi_{WN_Z|N_Z}^k$. This CF is taken as the approximation of $\bar{\varphi}^k$, i.e., $\bar{\varphi}_{k|k}^k$. Hence, for $k - N_Z + 1 \leq N_Z$ the FH initial condition in (79) approximates $\bar{\varphi}_{k|k}$, which is conditioned on the entire measurement history. Then, for $k - N_Z + 1 > N_Z$, the FH initial condition approximates the mean and variance of $\varphi_{WN_Z|N_Z}^{k-N_Z+1}$, produced by a previous iteration of this process.

$$\varphi_{WN_Z|N_Z}^k \approx \bar{\varphi}_{k-N_Z+1|k-N_Z+1}$$

Numerical comparisons have shown that the local initial condition found in this way performs well in reproducing the full information mean and variance. Moreover, simulations have shown that the finite horizon mean and variances agree very closely with the full information case even with horizon lengths as small as 8, as shown in the next section.

IX. Numerical Examples

We present a set of four examples demonstrating the performance of our proposed two-state estimator. The main challenge in implementing this estimator is the growth, with each measurement, of the number of terms needed to express the cpdf's CF. The proposed two-state estimator is more efficient and produces far fewer terms than the general-state estimator presented in [12, 13]. The improvement in performance is quantified in the table below, comparing the number of terms in the sum for a two-state implementation of [12, 13] to the number produced by the proposed algorithm, given by (68).

Measurement Update k	8	10	12
$N_{k\|k}$ of Previous [12, 13]	75036	1389207	25719609
$N_{k\|k}$ of Proposed	3193	21891	150049
Percent Retained	4.3%	1.6%	0.58%

However, the proposed estimator algorithm still suffers from the same fundamental issue of growing complexity. This motivated the use of a fixed window of the most recent measurements discussed in Section VIII, the performance of which is discussed next.

A. Finite Horizon Accuracy

Figure 1 shows, on a logarithmic scale, differences between the elements of the estimated state and error variance between the finite-horizon and full-information estimators, normalized to the full-information values; denote this normalized difference of a given element as $e(\cdot)$. The system parameters used are $\beta = 0.5$, $\gamma = 0.1$, $\alpha_1 = \alpha_2 = 0.8$, $\text{eig}(\Phi) = 0.8 \pm 0.55j$, $\mathrm{H} = [1 \quad 1]$, and $\Gamma = [0.5 \quad 1]^{\mathrm{T}}$. We compare the performance of horizon lengths of $N_Z = 8$ (dashed lines) and $N_Z = 10$ (solid lines). The subscripts indicate which element of the state estimate vector and error variance matrix are being compared. These results show that this finite horizon approximation is very accurate, with errors approximately between 0.01% and 0.0001% for our example and these two horizon lengths.

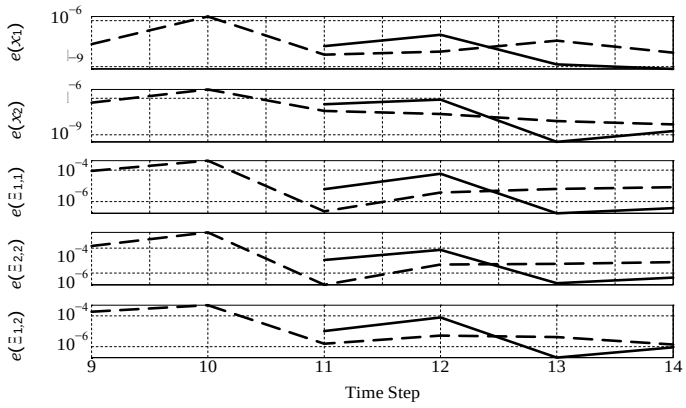


Fig. 1: Comparison of two finite horizon estimators (dashed is 8, solid is 10) to the full information estimator's means and error variances.

B. Cauchy-Gaussian Comparison

In our examples we compare the performance of our Cauchy estimator to the Kalman filter. To do this, we need to be able to choose Gaussian parameters for the Kalman filter that approximate the Cauchy parameters. To construct a normal or Gaussian pdf that best fits a given Cauchy pdf, the following optimization problem is solved

$$\sigma^* = \arg\min_{\sigma} \int_{-\infty}^{\infty} \left| f_X^C(x) - f_X^N(x) \right|^2 dx, \tag{80}$$

where the Cauchy pdf is $f_X^C(x) = \frac{\delta/\pi}{x^2+\delta^2}$, $\delta > 0$ and the normal pdf is given by $f_X^N(x) = \frac{e^{-x^2/(2\sigma^2)}}{\sqrt{2\pi\sigma}}$, $\sigma > 0$. Solving (80) analytically leads to a complex nonlinear equation relating σ^* to δ. Solving the latter numerically yields $\sigma^* = k_0\delta, \quad k_0 \approx 1.38980$.

C. Simulations

The simulations in Figs 2, 3, and 4 all use the same dynamics, where the eigenvalues of the transition matrix are $\text{eig}(\Phi) = 0.8 \pm 0.55j$, $\mathrm{H} = [1 \quad 1]$, and $\Gamma = [0.5 \quad 1]^{\mathrm{T}}$. The initial condition has a zero median and $\alpha_1 = \alpha_2 = 0.8$. All simulations use a measurement horizon length of $N_Z = 10$. In Fig. 2, $\gamma = 0.5$ and $\beta = 0.1$ so that the measurement noise dominates the process noise; in Fig. 3 the parameters are interchanged so that the process noise dominates the measurement noise. Gaussian parameters used for the LEG and for Gaussian noises are closest, in the L_2 sense, to their corresponding Cauchy distributions. For clarity of presentation in the figures, the first update occurs at $k = 0$ instead of $k = 1$.

Figures 2 and 3 compare the Cauchy and Kalman filters' responses to Cauchy distributed noises, and Fig. 4 compares their response to Gaussian distributed noises. Figures 2b and 3b show the same data as Figs. 2a and 3a when zoomed in around zero to demonstrate more clearly how the controllers respond when noise impulses are encountered.

When facing Cauchy distributed noises the proposed estimator outperforms the Kalman filter, especially when $\gamma > \beta$ as in Fig. 2. In this case, the Kalman filter's estimation error is almost always larger than that of the Cauchy estimator. Moreover, the impulsive noise values cause the conditional variance computed by the Kalman filter to be orders of magnitude smaller than the exact conditional variance computed by the Cauchy estimator.

In Fig. 3, where $\beta > \gamma$, both the Cauchy estimator's and the Kalman filter's estimation errors appear to have similar performance. However, the exact values of the conditional error variance computed by the Cauchy estimator are quite different from the error variance computed by the Kalman filter. In contrast to Fig. 2, in the case

where $\beta > \gamma$ the Kalman filter's estimation error decays faster due to the larger Kalman filter gain, since the process noise parameter dominates the measurement noise. This faster decay leads to the similar performance exhibited by both estimators.

During periods in the simulation without large impulses, the Cauchy and Kalman filters have similar performance, as shown in Figs. 2b and 3b. This suggests that in a non-impulsive noise setting, the two estimators would have similar performance. In a Gaussian noise setting, shown in Fig. 4, the Cauchy filter performs very well. It approximates the variance of the optimal Kalman filter and tracks its mean very closely. This demonstrates the robustness of the Cauchy estimator in a Gaussian noise environment.

X. Conclusions

An efficient two-state Cauchy estimation algorithm able to operate over an indefinite number of measurement updates is derived and presented. Although the estimator here is a special case of the general Cauchy estimation framework in [12, 13], the structures presented in this work take advantage of relationships currently understood only for the two-state system, allowing the development of an efficient recursive estimation structure. A method for using a finite window of measurement is proposed, and example simulations are presented to demonstrate the performance and robustness of the Cauchy estimator over a large number of measurement updates.

Appendix

Result 1: The product of two terms given by

$$\{a_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}v)\} \times \frac{1}{j\theta + \gamma + D \cdot \operatorname{sgn}(\mathrm{HA}v)} \cdot \frac{1}{j\theta - \gamma + D \cdot \operatorname{sgn}(\mathrm{HA}v)} = a_o^m + jd_o^m \cdot \operatorname{sgn}(\mathrm{HA}v) \quad (81a)$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ d_o^m \end{bmatrix} = \frac{1}{2\pi \cdot \bar{e}} \cdot \begin{bmatrix} \bar{a} & -\bar{d} \\ \bar{d} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ d_i^m \end{bmatrix}, \quad (81b)$$

and

$$\begin{aligned} \bar{a} &= -2\gamma, \quad D^2 - \gamma^2 - \theta^2, \\ \bar{d} &= 4\theta\gamma D, \\ \bar{e} &= \left(D^2 - \gamma^2 - \theta^2\right)^2 + 4(\theta D)^2 \end{aligned} \quad (81c)$$

Result 2: The product of two terms given by

$$\{a_i^m + b_i^m \cdot \operatorname{sgn}(\mathrm{HA}v) + jc_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}v)\} \times \frac{1}{j\theta + \gamma + \rho_m + D \cdot \operatorname{sgn}(\mathrm{HA}v)} \cdot \frac{1}{j\theta - \gamma + \rho_m + D \cdot \operatorname{sgn}(\mathrm{HA}v)} = a_o^m + b_o^m \cdot \operatorname{sgn}(\mathrm{HA}v) + jc_o^m + jd_o^m \cdot \operatorname{sgn}(\mathrm{HA}v), \quad (82a)$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ b_o^m \\ c_o^m \\ d_o^m \end{bmatrix} = \frac{1}{\Delta} \cdot \begin{bmatrix} \bar{e} & -\bar{f} & 0 & 0 \\ -\bar{f} & \bar{e} & 0 & 0 \\ 0 & 0 & \bar{e} & -\bar{f} \\ 0 & 0 & -\bar{f} & \bar{e} \end{bmatrix} \cdot \begin{bmatrix} \bar{a} & \bar{b} & -\bar{c} & -\bar{d} \\ \bar{b} & \bar{a} & -\bar{d} & -\bar{c} \\ \bar{c} & \bar{d} & \bar{a} & \bar{b} \\ \bar{d} & \bar{c} & \bar{b} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \quad (82b)$$

and

$$\begin{aligned} \bar{a} &= -2\gamma, \quad D^2 - \gamma^2 - \rho_m^2 - \theta^2, \\ \bar{c} &= 4\theta\gamma\rho_m, \\ \bar{e} &= \left(\rho_m^2 + D^2 - \gamma^2 - \theta^2\right)^2 + 4(\rho_m D)^2 + 4(\rho_m\theta)^2 + 4(\theta D)^2, \\ \Delta &= \pi \cdot \left(\bar{e}^2 - \bar{f}^2\right). \end{aligned} \quad (82c)$$

Result 3: The term given by

$$\begin{aligned} &\frac{a_i^m + b_i^m \cdot \operatorname{sgn}(\mathrm{HA}v) + jc_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}v)}{j\theta + \rho_m - \gamma \cdot \operatorname{sgn}\left(B_{k|k}^m v\right) + D \cdot \operatorname{sgn}(\mathrm{HA}v)} \\ &\quad - \frac{a_i^m - b_i^m \cdot \operatorname{sgn}(\mathrm{HA}v) - jc_i^m + jd_i^m \cdot \operatorname{sgn}(\mathrm{HA}v)}{j\theta - \rho_m - \gamma \cdot \operatorname{sgn}\left(B_{k|k}^m v\right) + D \cdot \operatorname{sgn}(\mathrm{HA}v)} \\ &= a_o^m + b_o^m \cdot \operatorname{sgn}\left(B_{k|k}^m v\right) \operatorname{sgn}(\mathrm{HA}v) + jd_o^m \cdot \operatorname{sgn}(\mathrm{HA}v) \\ &\quad + jc_o \cdot \operatorname{sgn}\left(B_{k|k}^m v\right) \end{aligned} \quad (83a)$$

where the i subscript denotes the input term and the o subscript denotes the output term,

$$\begin{bmatrix} a_o^m \\ b_o^m \\ c_o^m \\ d_o^m \end{bmatrix} = \frac{1}{\Delta} \cdot \begin{bmatrix} \bar{e} & -\bar{f} & 0 & 0 \\ -\bar{f} & \bar{e} & 0 & 0 \\ 0 & 0 & \bar{e} & -\bar{f} \\ 0 & 0 & -\bar{f} & \bar{e} \end{bmatrix} \cdot \begin{bmatrix} \bar{a} & \bar{b} & -\bar{c} & -\bar{d} \\ \bar{b} & \bar{a} & -\bar{d} & -\bar{c} \\ \bar{c} & \bar{d} & \bar{a} & \bar{b} \\ \bar{d} & \bar{c} & \bar{b} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} -\rho_m & D & -\theta & 0 \\ 0 & -\gamma & 0 & 0 \\ 0 & 0 & -\gamma & 0 \\ 0 & \theta & D & -\rho_m \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \quad (83b)$$

and

$$\begin{aligned} \bar{a} &= \gamma_2 + D_2 - \rho_m - \theta_2, \quad \bar{b} = -2\gamma D, \\ \bar{c} &= -2\theta\gamma, \quad \bar{d} = 2\theta D, \\ \bar{e} &= \bar{a}^2 + \bar{b}^2 + \bar{c}^2 + \bar{d}^2, \quad \bar{f} = 2\left(\bar{a}\bar{b} + \bar{c}\bar{d}\right), \\ \Delta &= \pi \cdot \left(\bar{e}^2 - \bar{f}^2\right). \end{aligned} \quad (83c)$$

References

[1] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

[2] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

[3] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution," *IEEE Communications Letters*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[4] P. Reeves, "A non-gaussian turbulence simulation," Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-69-67, 1969.

[5] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall, 1994.

[6] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[7] G. A. Tsihrintzis, *Statistical Modeling and Receiver Design for Multi-User Communication Networks;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[8] G. A. Hewer, R. D. Martin, and J. Zeh, "Robust preprocessing for kalman filtering of glint noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 120–128, January 1987.

[9] J. P. Nolan, J. G. Gonzalez, and R. C. Núñez, "Stable filters: A robust signal processing framework for heavy-tailed noise." in *In Proceedings of the 2010 IEEE Radar Conference*, 2010.

[10] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems,"

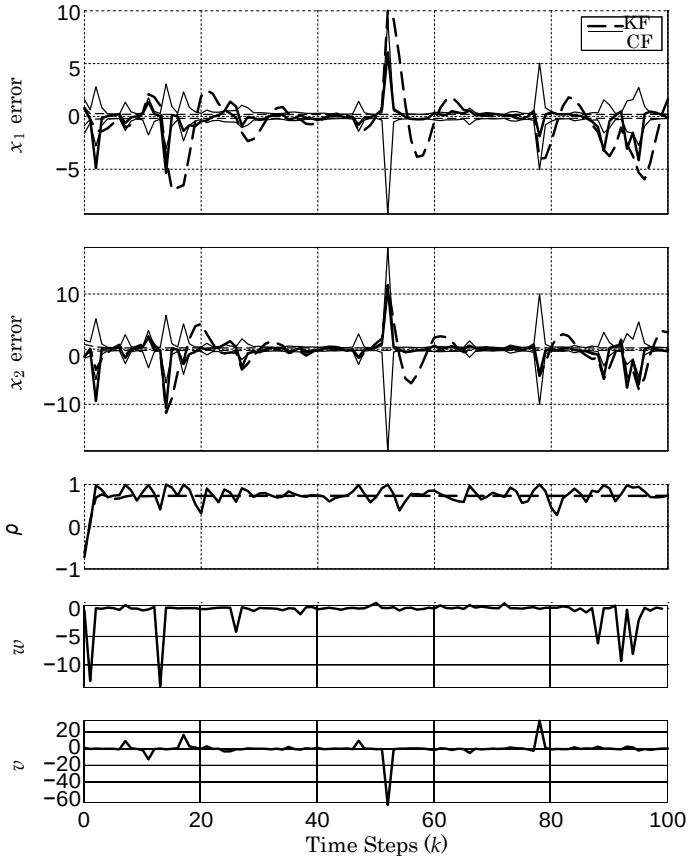
□ ¯ ¯ □ □ □
$\bar{d}$ $\bar{c}$ $\bar{a}$ m
$\bar{b}$

IEEE Transactions on Automatic Control, vol. 55, no. 6, pp. 1329–1342,

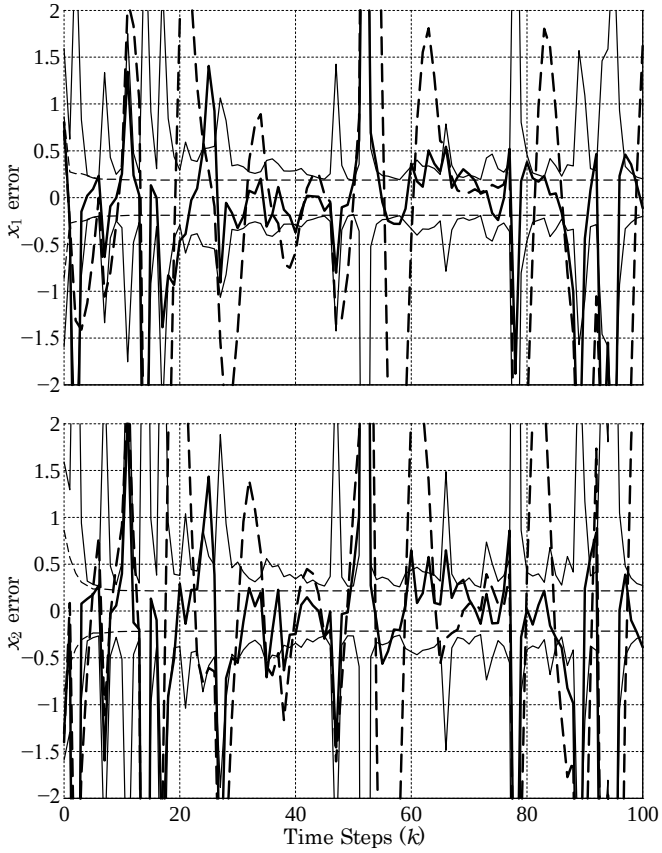
d_i

(82b) 2010.

[11] ——, "State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach," *SIAM J. Control Optim.*, vol. 50, no. 4, pp. 1971–1994, 2012.
[12] ——, "Multivariate Cauchy Estimator with Scalar Measurement and Process Noises," in *Proceedings of the* 52nd *IEEE Conference on Decision and Control*, Florence, Italy, December 2013.
[13] ——, "Multivariate cauchy estimator with scalar measurement and process noises," *Submitted to SIAM*, 2013.
[14] ——, "Characteristic Function Approach for Estimation of Scalar Systems with Cauchy Noises," in *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario Canada, Aug. 2010.
[15] D. Champeney, *A Handbook of Fourier Theorems*. New York: Cambridge University Press, 1987.
[16] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. New York: Oxford University Press, 1999.

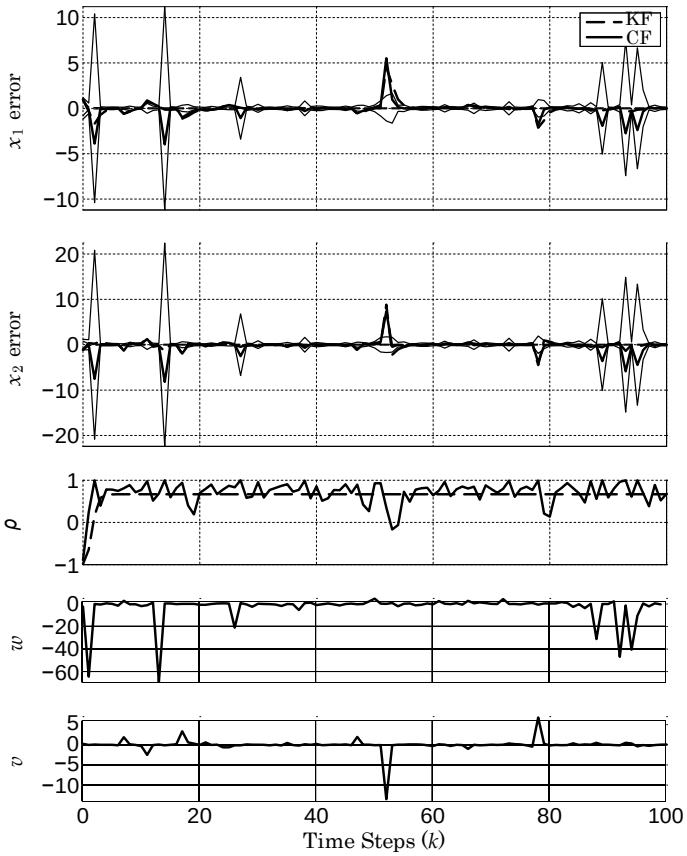


(a) Full view.

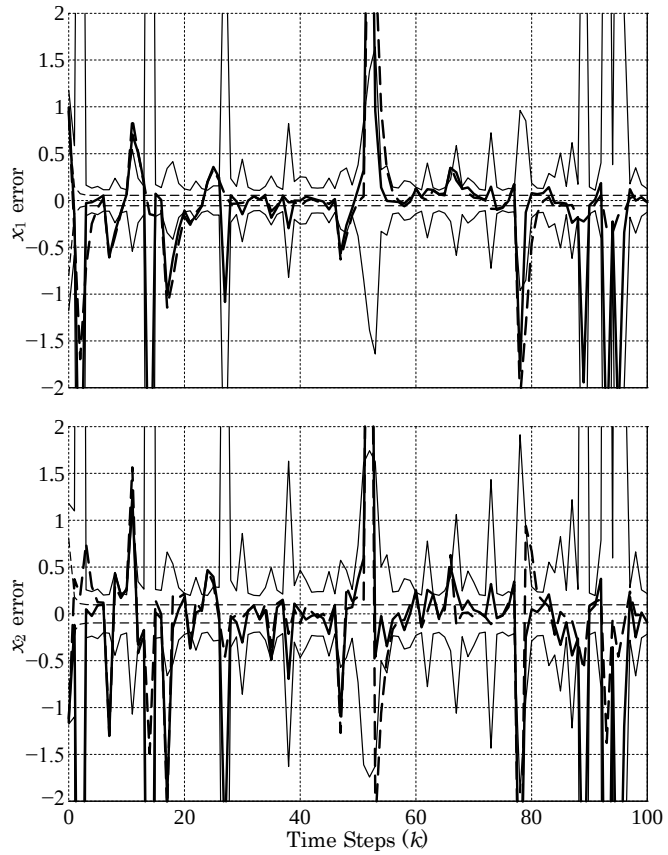


(b) Zoomed view.

Fig. 2: Cauchy and Kalman estimators for $\gamma > \beta$; thick lines are the estimate errors, and thin lines are the standard deviations.

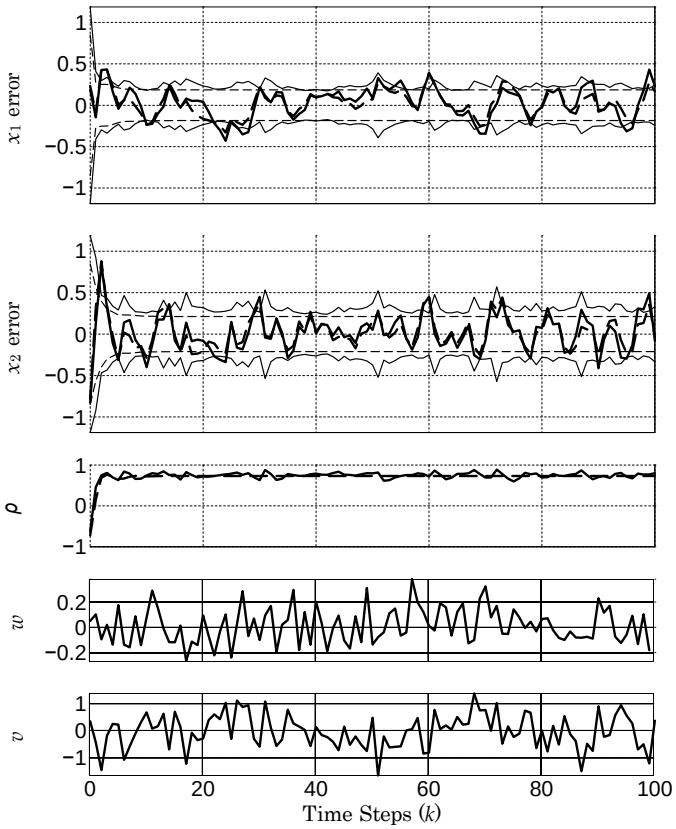


(a) Full view.

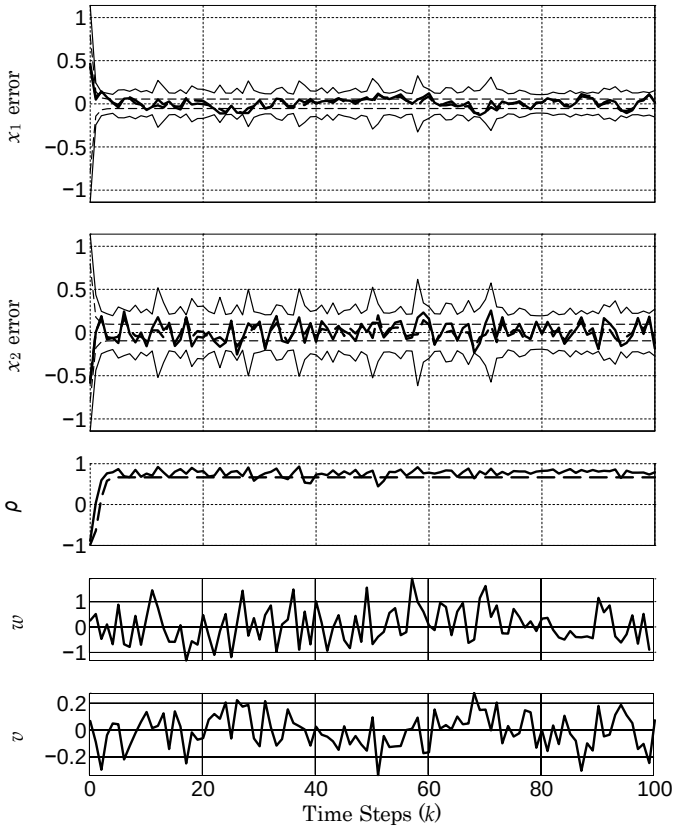


(b) Zoomed view.

Fig. 3: Cauchy and Kalman estimators for $\beta > \gamma$; thick lines are the estimate errors, and thin lines are the standard deviations.



(a) Gaussian noises, $\beta = 0.1$ and $\gamma = 0.5$.



(b) Gaussian noises, $\beta = 0.5$ and $\gamma = 0.1$.

Fig. 4: Cauchy and Kalman estimators against Gaussian noises; thick lines are the estimate errors, and thin lines are the standard deviations.

D Appendix

Jason L. Speyer, Moshe Idan, and Javier H. Fernández
A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise
Automatica, to appear

Automatica xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A stochastic controller for a scalar linear system with additive Cauchy noise☆

Jason L. Speyer[a,1], Moshe Idan[b], Javier H. Fernández[a]

[a] Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095-1597, United States

[b] Aerospace Engineering, Technion - Israel Institute of Technology, Haifa, 32000, Israel

article info

Article history:
Received 1 May 2012
Received in revised form
13 June 2013
Accepted 6 October 2013
Available online xxxx

Keywords:
Stochastic control
Optimal controller synthesis for systems with uncertainties
Heavy tailed distributions

abstract

A stochastic control scheme is developed for scalar, discrete-time, and linear-dynamic systems driven by Cauchy distributed process and measurement noises. When addressing the optimal control problem for such systems, the standard quadratic cost criteria cannot be used. In this study we introduce a new objective function that is functionally similar to the Cauchy probability density function. The performance index, defined as the expectation of this objective function with respect to the Cauchy densities, exists. The dynamic programming solution to the fixed and finite horizon optimal control problem that uses this performance index appears to be intractable. Therefore, a moving horizon optimal model predictive control problem is implemented, for which the conditional expected value of the objective function and its gradients can be computed in *closed* form and without assumptions such as certainty equivalence. Numerical results are shown for this m-step model predictive optimal controller and compared to a similar, Linear-Exponential-Gaussian model predictive controller. An essential difference between the Cauchy and Gaussian controllers when applied to a system with Cauchy noises is that, while the Gaussian controller is linear and reacts strongly to all noise pulses, the Cauchy controller can differentiate between measurement and process noise pulses by ignoring the former while responding to the latter. This property of the Cauchy controller occurs when an impulsive measurement noise is more likely than an impulsive process noise. The Cauchy and Gaussian controllers react similarly when applied to a system with Gaussian noises, demonstrating the robustness of the proposed control scheme.

1. Introduction

Modern stochastic optimal control algorithms, such as the Linear-Quadratic-Gaussian (LQG) and Linear-Exponential-Gaussian (LEG) algorithms, assume that the system is driven by additive Gaussian process and measurement noises (Speyer & Chung, 2008, Chapters 9 and 10). Since the Gaussian probability density function (pdf) is a light-tailed pdf, which essentially rules out the possibility of large deviations, these algorithms are unable to handle measurement outliers that produce large filter residuals due to impulsive changes in the measurements. For example, the governing types of noises that occur in radar and sonar applications are atmospheric and underwater acoustic noises. Noises of this type exhibit very impulsive behaviors that are not captured by a Gaussian distribution (Kuruoglu, Fitzgerald, & Rayner, 1998). Dynamic impulsive noises can be used in modeling unknown adversarial motion, as well as for modeling air turbulence, which was shown to be better represented by non-Gaussian, heavy-tailed distributions (Reeves, 1969). The literature on how to handle outliers is dominated by heuristic methods that assume a Gaussian distribution for the underlying stochastic processes. Moreover, these methods typically work for *a posteriori* analysis of static problems, which is inadequate for control (Fernholz, Morgenthaler, & Tukey, 2004; Hampel, Ronchetti, Rousseeuw, & Stahel, 1968; Holland & Welsch, 1977; Pirinen, 2008).

In this paper, the proposed controller is based on a discrete time linear dynamic system with Cauchy distributed process and measurement noises and initial condition. The Cauchy probability distribution function (pdf) is in the class of probability distributions called symmetric alpha-stable ($S\alpha$-S) distributions, whose members are described using their characteristic functions (see Samorodnitsky & Taqqu, 1994 for a comprehensive treatment of $S\alpha$-S densities). Within this class: the Gaussian pdf corresponds to $\alpha = 2$; $\alpha = 1$ leads to the Cauchy pdf; all pdf's with $\alpha < 2$ have

☆ This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-10-1-0570, and by the United States–Israel Binational Science Foundation, Grant 2008040. The material in this paper was partially presented at the IEEE Conference on Decision and Control (CDC), December 15–17, 2010, Atlanta, Georgia, USA. This paper was recommended for publication in revised form by Associate Editor Gang Tao under the direction of Editor Miroslav Krstic.

E-mail addresses: speyer@seas.ucla.edu (J.L. Speyer), moshe.idan@technion.ac.il (M. Idan), jhf@seas.ucla.edu (J.H. Fernández).

[1] Tel.: +1 310 206 4451; fax: +1 310 206 2302.

0005-1098/$ – see front matter © 2013 Published by Elsevier Ltd
http://dx.doi.org/10.1016/j.automatica.2013.11.005

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

an infinite variance;[2] and the Cauchy pdf lacks a defined mean or first moment. However, for Cauchy uncertainties the conditional mean of the state and its conditional variance given a measurement history do exist (Idan & Speyer, 2008, 2010). An important aspect as well is that the conditional error variance is a function of the measurement history (Idan & Speyer, 2008, 2010), quite *unlike* the Gaussian case where the conditional error variance is *a priori* known.

One of the fundamental lessons learned from the results presented in this paper is that the handling of measurement outliers directly involves the stochastic controller. A stochastic control law has to be consistent with the underlying structure of the conditional pdf of the state variable and with the assumptions used in its construction. Under the Gaussian assumption, only the state estimate depends on the measurement while the error variance of the estimate does not. Since optimal controllers based on Gaussian noises (e.g. LQG and LEG controllers) are linear in the measurements, they do not differentiate between outliers and normal measurements and respond to both of them in the same manner.

Since stochastic optimal control algorithms are developed based on the minimization or maximization of an expectation of an objective function, an appropriate computable objective function has to be determined for systems with Cauchy noises. The objective functions normally used in the LQG and the positive LEG (by which we mean that the quadratic argument of the exponential is positive) control settings are not suitable here because those expectations are either undefined or infinite for Cauchy pdf's.[3] Hence, a different objective function has to be chosen for this case. We chose an objective function that is a product of functions that resemble scaled Cauchy pdf's in structure and depend on the state or control variable. For Cauchy uncertainty, the expectation of this objective function is finite and its conditional expectation given the measurement history can be expressed as a *closed form* function of the measurements and controls. A similar performance index formulation was made for the LEG problem, where the objective function was constructed as a product of exponential functions of the state and the control (Jacobson, 1973) resembling Gaussian pdf's in form.

Based on the proposed objective function, a dynamic programming recursion rule is developed to construct the optimal control function. However, the application of this recursion rule to the finite fixed horizon optimal control problem for systems with Cauchy distributed noises appears to be intractable. Therefore, we instead develop an m-step model predictive optimal controller (Morari & Zafiriou, 1989), i.e., open-loop-optimal feedback controller,[4] for such systems. Some initial results on scalar Cauchy model predictive control, the Cauchy controller, were presented in Idan, Emadzadeh, and Speyer (2010) and Speyer, Idan, and Fernández (2010). The current work presents a more complete scalar Cauchy m-step optimal model predictive solution. Although formulating an m-step optimal predictive controller produces a suboptimal solution, the performance index is evaluated by determining the conditional expected value of the objective function given the measurement history exactly in closed form. This Cauchy controller is expressed as a nonlinear function of the measurement history.

The scalar state problem addressed in this work provides insight into the problem of handling outliers, which is resolved by the Cauchy stochastic controller explicitly, and not in the filter alone, as has been traditionally done for Gaussian noises. There is no equivalent construction of such a Cauchy controller in the non-Gaussian stochastic control literature, including in the stochastic model predictive control (MPC) setting. In stochastic MPC, the objective is to minimize a quadratic objective function subject to probabilistic inequality constraints. In Jun and Bitmead (2005) the assumption of a linear system with additive Gaussian noise is essential in transforming the stochastic optimal control problem into a deterministic one.[5] In Cannon, Cheng, and Rakovic (2012) the Gaussian assumption is removed by assuming that all stochastic uncertainties have bounded support. Our motivation is to determine a stochastic controller based on Cauchy pdf's with infinite support, which produces a deterministic analytic performance index in the control and measurement history.

The Cauchy controller gives insight into robustness by introducing heavy tailed distributions into the design. Moreover, a controller design process is advocated for handling outliers based on whether the measurement noise dominates the process noise. If measurement noise dominates, the measurement outliers do not have to be known, since the controller responds little to them.

The paper is organized as follows. The optimal control problem is formulated in Section 2. A new objective function is introduced in Section 3 having the functional form of the Cauchy pdf. The performance index to be optimized is the conditional expectation of the objective function. A dynamic programming recursion rule is derived in Section 4. Since the solution to the dynamic programming problem appears intractable, the Cauchy m-step model predictive optimal controller is formulated and developed in Section 5. The conditional expected value of the objective function is determined in closed form. However, due to its complexity, the maximization of the conditional performance index with respect to the projected control sequence is determined numerically, as detailed at the end of Section 5. Numerical examples are presented in Section 6. First, a one-step process with a single measurement is first explored to gain insight into the structure and behavior of the Cauchy controller in the presence of impulsive heavy-tailed noises discussed in Section 6.1. Then, multi-step examples are given in Section 6.2 comparing the performance of the Cauchy and Gaussian m-step model predictive optimal controllers under both Cauchy and Gaussian noises. Finally, concluding remarks are presented in Section 7.

2. Problem statement

Consider the linear, discrete-time, scalar stochastic system

$$\mathbf{x}_k = \Phi\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}, \qquad \mathbf{z}_k = H\mathbf{x}_k + \mathbf{v}_k, \tag{1}$$

where $\mathbf{x}_k$ is the state, $\mathbf{u}_k$ is the control signal, $\mathbf{z}_k$ is the measurement, and k is the time index. The signals $\mathbf{w}_k$ and $\mathbf{v}_k$ are independent process and measurement noise sequences, respectively, that are assumed to be independent of each other and Cauchy distributed with pdf's

$$f_{\mathbf{w}_k}(w_k) = \frac{\beta/\pi}{w_k^2 + \beta^2}, \qquad f_{\mathbf{v}_k}(v_k) = \frac{\gamma/\pi}{v_k^2 + \gamma^2}. \tag{2}$$

[2] The in-phase component of radar clutter time series agrees extremely well with a $S\alpha S$ pdf with $\alpha = 1.7$ (Tsakalides & Nikias, 1998).

[3] In this paper, we will compare our Cauchy controller to the negative LEG controller.

[4] Open-loop-optimal feedback was introduced in the stochastic control context by Dreyfus in Dreyfus (1965)

[5] Moreover, the solution in Jun and Bitmead (2005) is exact only for the scalar case.

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

These pdf's have a median of zero and scale parameters $\beta > 0$ and $\gamma > 0$, respectively. The initial state is assumed to be also Cauchy distributed

$$f_{\mathbf{x}_0}(x_0) = \frac{\alpha/\pi}{(x^0 - \bar{x}_0)^2 + \alpha^2} \tag{3}$$

with a median at $\bar{x}_0$ and scale parameters $\alpha > 0$.

Define the measurement history as

$$Z_k := \{\mathbf{z}_0, \ldots, \mathbf{z}_k\}, \tag{4}$$

and assume that $\mathbf{u}_k \in \mathsf{F}_k$, where F_k is the class of functions adapted to the σ-algebra σ_k generated by Z_k. The notion of adaptation is that the random variable, here the control, is measurable with respect to events in σ_k.

In this work we consider a regulation problem, where the state is to be driven to zero while penalizing the control effort. This problem will be cast as an optimal control problem, the performance index of which will be introduced next.

3. Performance index for a Cauchy controller

In posing an optimal control problem for the model in (1), commonly used objective functions like the quadratic or the exponential of a quadratic cannot be used because the expectations required to evaluate those objective functions are infinite when the system noise inputs have heavy-tailed Cauchy pdf's. Therefore, one has to introduce a new, computable objective function. In this work we suggest an objective function that resembles in its form the Cauchy pdf, and which also allows an analytical derivation of the controller. This general objective function is reminiscent to the choice of the objective function for the LEG (Speyer & Chung, 2008, Chapter 10), which was constructed as a product of functions resembling the Gaussian pdf. The original motivation for the LEG objective function (Jacobson, 1973) was to consider these exponential functions as membership functions in fuzzy set theory, where the objective function was constructed as a product of these Gaussian-shaped membership functions. Similarly, a objective function constructed from products of Cauchy-shaped membership functions is proposed here for systems with Cauchy noises. This similarity allows us to compare the optimal controllers for linear systems with analogous performance index and different types of noise.

Consequently, the membership functions that penalize the state and control are chosen as rational functions resembling Cauchy pdf's and are expressed as

$$M_x(\mathbf{x}_k) = \frac{\eta_k^2}{\mathbf{x}_k^2 + \eta_k^2}, \qquad M_u(\mathbf{u}_k) = \frac{\zeta_k^2}{\mathbf{u}_k^2 + \zeta_k^2}, \tag{5}$$

where η_k and ζ_k are design parameters which affect the shape of these functions. Note that as $\eta_k \to \infty$ or $\zeta_k \to \infty$, the penalties on the state or the control are removed. In addition, smaller values of η_k and ζ_k induce heavier weighings on the respective variables. These particular functions are chosen because they make the expectation with respect to the conditional pdf generated by Cauchy noise of the resulting performance index analytic in the control and measurement history as will be shown in the sequel.

Ideally, for regulation, we would like to solve an infinite horizon regulation problem. This is often done by solving a fixed horizon problem from $k = 0$ to N, e.g., using the dynamic programming approach, and then letting N go to infinity. However, we define a more general objective function that will serve both the dynamic programming formulation and the model predictive objective function, discussed later, as

$$\psi\left(X_{\ell+1}^{n}, U_{\ell}^{n-1}\right) = \prod_{i=\ell}^{n-1} M_x(\mathbf{x}_{i+1}) M_u(\mathbf{u}_i) = \prod_{i=\ell}^{n-1} \frac{\eta_{i+1}^2}{\mathbf{x}_{i+1}^2 + \eta_{i+1}^2} \cdot \frac{\zeta_i^2}{\mathbf{u}_i^2 + \zeta_i^2} \tag{6}$$

with $\ell < n$. The state and control histories used in the objective function are defined as

$$X_{\ell}^{n} := \{\mathbf{x}_\ell, \ldots, \mathbf{x}_n\}, \tag{7a}$$

$$U_{\ell}^{n} := \{\mathbf{u}_\ell, \ldots, \mathbf{u}_n\}. \tag{7b}$$

Since the performance index is composed of the expected value of bell-shaped curves in (6), we maximize to obtain regulation of the state with penalties on the control. The performance index, which is the expectation of the objective function in (6), is given by

$$J_{\ell,n}^{*} = \max_{U_{\ell}^{n-1}} E\left\{\psi\left(X_{\ell+1}^{n}, U_{\ell}^{n-1}\right)\right\}, \tag{8}$$

where element $\mathbf{u}_k$ in U_{ℓ}^{n-1}, the control sequence, is $\mathbf{u}_k \in \mathsf{F}_k$. The expectation is taken over all the underlying random variables, i. e., all the measurement and process noises, and the initial conditions.

4. Dynamic programming approach

To solve the problem in (8), we first consider the dynamic programming approach with the hope of constructing an analytic form of the controller. It requires the derivation of the general optimal return function and the backwards recursion rule, which are derived next. This derivation shows the complexity of evaluating the backwards recursion rule for Cauchy noise patterns and motivates the derivation of an alternative model-predictive-type controller.

In this derivation, we consider a given fixed horizon problem by setting ℓ and n of (8) to be $\ell = 0$ and $n = N$. Expanding the expectation on the right-hand side of (8) results in

$$\begin{aligned}
J_{0,N}^{*} &= \max_{U_0^{N-1}} E\left\{\psi\left(X_1^N, U_0^{N-1}\right)\right\} \\
&= \max_{U_0^{N-1}} E\left\{E\left\{\psi\left(X_1^N, U_0^{N-1}\right) \middle| Z_N\right\}\right\} \\
&= \max_{U_0^{N-1}} E\left\{E\left\{\cdots E\left\{E\left\{\psi\left(X_1^N, U_0^{N-1}\right) \middle| Z_N\right\} \middle| Z_{N-1}\right\} \cdots \middle| Z_0\right\}\right\} \\
&= E\left\{\max_{\mathbf{u}_0 \in \mathsf{F}_0} E\left\{\cdots \max_{\mathbf{u}_{N-1} \in \mathsf{F}_{N-1}} E\left\{E\left\{\psi\left(X_1^N, U_0^{N-1}\right) \middle| Z_N\right\} \middle| Z_{N-1}\right\} \cdots \middle| Z_0\right\}\right\} \\
&= \int_{-\infty}^{\infty} \max_{\mathbf{u}_0 \in \mathsf{F}_0} \int_{-\infty}^{\infty} \cdots \max_{\mathbf{u}_{N-1} \in \mathsf{F}_{N-1}} \int_{-\infty}^{\infty} E\left\{\psi\left(X_1^N, U_0^{N-1}\right) \middle| Z_N\right\} f_{\mathbf{z}_N|\sigma_{N-1}}(z_N | Z_{N-1})\, dz_N \\
&\quad \cdots f_{\mathbf{z}_1|\sigma_0}(z_1 | Z_0)\, dz_1 \; f_{\mathbf{z}_0}(z_0)\, dz_0,
\end{aligned} \tag{9}$$

where element $\mathbf{u}_k$ in U_0^{n-1} is $\mathbf{u}_k \in \mathsf{F}_k$, and $f_{\mathbf{z}_k|\sigma_{k-1}}(z_k | Z_{k-1})$ is the probability density function conditioned on the σ_{k-1} generated by the measurement histories up to $k-1$. The Fundamental Lemma (Speyer & Chung, 2008, Chapter 9) was used to interchange the maximization and expectation operations in the third equality of (9).

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

To apply dynamic programming to the problem in (9), we need to recast the performance index as a recursion rule for an optimal return function. To get this recursion, first note that the optimal performance index starting at time step $k+1$, which is embedded in (9), is

$$\max_{U_{k+1}^{N-1}} E\{\psi(X_1^N, U_0^{N-1}) \mid Z_{k+1}\} = \max_{\mathbf{u}_{k+1}\in F_{k+1}} \int_{-\infty}^{\infty} \cdots \max_{\mathbf{u}_{N-1}\in F_{N-1}} \int_{-\infty}^{\infty} E\{\psi(X_1^N, U_0^{N-1}) \mid Z_N\} f_{\mathbf{z}_N|\sigma_{N-1}}(z_N|Z_{N-1})\, dz_N \cdots f_{\mathbf{z}_{k+2}|\sigma_{k+1}}(z_{k+2}|Z_{k+1})\, dz_{k+2}. \quad (10)$$

Using the above, the performance index (9) can be restated as

$$J_{0,N}^* = \max_{\mathbf{u}_0\in F_0} \int_{-\infty}^{\infty} \cdots \max_{\mathbf{u}_k\in F_k} \int_{-\infty}^{\infty} \max_{U_{k+1}^{N-1}} E\{\psi(X_1^N, U_0^{N-1}) \mid Z_{k+1}\} f_{\mathbf{z}_{k+1}|\sigma_k}(z_{k+1}|Z_k) \cdots f_{\mathbf{z}_1|\sigma_0}(z_1|Z_0) f_{\mathbf{z}_0}(z_0)\, dz_{k+1} \cdots dz_1\, dz_0. \quad (11)$$

To construct the dynamic programming solution, define the optimal return function as

$$\bar{J}_{k+1,N}^*(Z_{k+1}) \text{ t:. } \max_{U_{k+1}^{N-1}} E\{\psi(X_1^N, U_0^{N-1}) \mid Z_{k+1}\} \times f_{\mathbf{z}_{k+1}|\sigma_k}(z_{k+1}|Z_k) \cdots f_{\mathbf{z}_1|\sigma_0}(z_1|Z_0) f_{\mathbf{z}_0}(z_0), \quad (12)$$

which is the integrand of the most inner integral in (11). Then, the sought after dynamic programming recursion rule becomes

$$\bar{J}_{k,N}^*(Z_k) = \max_{\mathbf{u}_k\in F_k} \int_{-\infty}^{\infty} \bar{J}_{k+1,N}^*(Z_{k+1})\, dz_{k+1}. \quad (13)$$

This is the general form for the recursion rule from which the optimal performance index is determined by

$$J_{0,N}^* = \int_{-\infty}^{\infty} \bar{J}_{0,N}^*(Z_0)\, dz_0. \quad (14)$$

A method for obtaining the conditional pdf's needed to evaluate the expectations in (12) is presented in Idan and Speyer (2010) and reviewed in Appendix A. It can be verified that the functional dependence of these conditional pdf's on the measurements is very complex, making the integration and maximization required in (13) intractable.[6] The main difficulty arises from the need to average over future measurements.

Alternatively, a model predictive controller that does not average over future measurements and can therefore be constructed more feasibly. Although sub-optimal, it will produce a practical and computable control solution for a system with Cauchy noises. This paper considers an m-step horizon optimal model predictive controller, which is developed in the next section.

[6] Note that for a Gaussian system with a multiplicative objective function constructed from the product of exponentials with quadratic arguments, (13) produces a linear controller — the LEG controller (Speyer & Chung, 2008, Chapter 10). However, if the objective function is constructed from a sum of exponentials with quadratic arguments, the solution to (13) appears intractable.

5. Cauchy optimal model predictive controller

The Cauchy m-step optimal model predictive controller is determined by maximizing a moving, fixed-horizon performance index that is a function of all the measurements up to the current time but no future ones. The performance index is defined as the expected value of the objective function in (6) using $\ell = k$ (i.e., the current time) and $n = k + m$ t:. p, where m is the size of a moving horizon. The expectation is carried out over the stochastic variables associated with the states in this moving horizon, i.e., the future process noise, and the current measurement history. As opposed to dynamic programming, here we define the predictive control sequence as U_k^{p-1}, where every element in it is in F_k; i.e. $\mathbf{u}_\ell \in F_k \ \forall\ \ell = k, \ldots, p-1$. Hence, the performance index is given as

$$J_{k,p}^* = \max_{U_k^{p-1}} E\{\psi(X_{k+1}^p, U_k^{p-1}) \mid Z_k\}, \quad (15)$$

where $E[\cdot]$ is the expectation taken over Z_k and X_{k+1}^p. It is important to note that since $\psi(\cdot,\cdot)$ of (6) is bounded by one for all X_{k+1}^p and U_k^{p-1}, the performance index $J_{k,p} = E\{\psi(X_{k+1}^p, U_k^{p-1}) \mid Z_k\}$ of (15) always exists and is finite.

The optimal performance index of (15) is restated as

$$J_{k,p}^* = \max_{U_k^{p-1}} E\{E\{\psi(X_{k+1}^p, U_k^{p-1}) \mid Z^k\} \mid Z_k\} = E\{\max_{U_k^{p-1}} E\{\psi(X_{k+1}^p, U_k^{p-1}) \mid Z_k\} \mid Z_k\} \text{ t:. } E\{J_{Z_k}^*\}, \quad (16)$$

where the Fundamental Lemma (Speyer & Chung, 2008, Chapter 9) is used to interchange the maximization and expectation operations. The current control $\mathbf{u}_k$ is found as a function of the current information pattern that includes both current and past measurements and the past control inputs. An important characteristic of the objective function $\psi(\cdot,\cdot)$ is that for the Cauchy densities the conditional performance index J_{Z_k} of (16) can be determined in a closed form and the optimal control signal can be determined by maximizing this expression.

Up until now we have considered that the control $\mathbf{u}_k$ is adapted to the σ-algebra σ_k generated by the measurement history, Z_k. However, if the state is decomposed into a dynamic system that contains all the underlying random variables and another dynamic system only driven by the control, then the determination of the optimal control is not only simplified, but indeed tractable. Given this decomposition, we show that the control has to be adaptive to only the σ-algebra generated by the measurement history of the decomposed state associated with the underlying random variables.

Consider the linear, discrete-time, scalar stochastic system of (1) with the measurement history given by (4). Let $\mathbf{u}_k$ be adaptive to the filtration σ-algebra σ_k generated by the measurement history Z_k. Filtration implies that the collection of σ-algebras σ_k have the property that if $j \le k$, then $\sigma_j \subseteq \sigma_k$ (Fleming & Rishel, 1975). Therefore, filtration is the evolution of the σ-algebra generated by measurement history through time. Adaptation means that the control is a measurable function of events on this σ-algebra, i.e., this ensures that the control sequence is causal. Now consider the decomposition $\mathbf{x}_k = \tilde{\mathbf{x}}_k + \bar{\mathbf{x}}_k$ where

$$\tilde{\mathbf{x}}_k = \Phi\tilde{\mathbf{x}}_{k-1} + \mathbf{w}_{k-1}, \qquad \tilde{\mathbf{z}}_k = H\tilde{\mathbf{x}}_k + \mathbf{v}_k, \quad (17a)$$

$$\bar{\mathbf{x}}_k = \Phi\bar{\mathbf{x}}_{k-1} + \mathbf{u}_{k-1}, \qquad \bar{\mathbf{z}}_k = H\bar{\mathbf{x}}_k. \quad (17b)$$

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

Here, $\tilde{\mathbf{x}}_k$ and $\tilde{\mathbf{z}}_k$ are the state and the measurement of the subsystem containing all the underlying random variables, i.e., $\mathbf{w}_k$, $\mathbf{v}_k$, and the initial condition $\mathbf{x}_0$, which is Cauchy distributed with zero median similar to (3). Similarly, $\bar{\mathbf{x}}_k$ and $\bar{\mathbf{z}}_k$ are the state and measurement of a dynamic system driven by $\mathbf{u}_k$ with initial condition $\bar{\mathbf{x}}_0$. The initialization of either system is arbitrary and does not affect the simplicity introduced by the decomposition, as long as the sum of the initial value of (17b) and the median of $\tilde{\mathbf{x}}_0$ equals the given $\bar{\mathbf{x}}_0$.

The measurement history can be decomposed as $Z_k = \tilde{Z}_k + \bar{Z}_k$ where

$$\tilde{Z}_k = \{\tilde{\mathbf{z}}_0, \ldots, \tilde{\mathbf{z}}_k\}, \qquad \bar{Z}_k = \{\bar{\mathbf{z}}_0, \ldots, \bar{\mathbf{z}}_k\}. \tag{18}$$

In the following it is shown that the control is measurable on events generated by $\tilde{Z}_k$ only.

Theorem 1. *Consider the filtration σ-algebra $\tilde{\sigma}_k$ generated by $\tilde{Z}_k$, with the decomposition $Z_k = \tilde{Z}_k + \bar{Z}_k$. For $\tilde{Z}_k \in \tilde{\sigma}_k$ and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, $\bar{Z}_k$ is adapted to $\tilde{\sigma}_{k-1}$ and $\mathbf{u}_k$ is adapted to $\tilde{\sigma}_k$.*

Proof. Start with $k = 0$. The initial state is decomposed as $\mathbf{x}_0 = \tilde{\mathbf{x}}_0 + \bar{\mathbf{x}}_0$, where $\bar{\mathbf{x}}_0$ is a given non-random parameter. The measurement decomposes as $\mathbf{z}_0 = \tilde{\mathbf{z}}_0 + \bar{\mathbf{z}}_0$, where $\bar{\mathbf{z}}_0 = H\bar{\mathbf{x}}_0$ is a given non-random parameter and $\tilde{\mathbf{z}}_0 = \tilde{Z}_0 \in \tilde{\sigma}_0$. Then, $\mathbf{u}_0$, which is determined by $\mathbf{z}_0$, is adapted to $\tilde{\sigma}_0$. At $k = 1$, both $\bar{\mathbf{x}}_1 = \Phi\bar{\mathbf{x}}_0 + \mathbf{u}_0$ and $\bar{\mathbf{z}}_1 = H\bar{\mathbf{x}}_1$ are adapted to $\tilde{\sigma}_0$, and thus $\bar{Z}_1$ is adapted to $\tilde{\sigma}_0$. For the measurement at $k = 1$, $\tilde{\mathbf{z}}_1 \in \tilde{\sigma}_1$, $\tilde{Z}_1 \in \tilde{\sigma}_1$, and $\tilde{\sigma}_0 \subset \tilde{\sigma}_1$. Hence, since $\mathbf{u}_1$ is determined by $Z_1 = \tilde{Z}_1 + \bar{Z}_1$, it is adapted to $\tilde{\sigma}_1$. Recursively to any k, $\bar{Z}_k$ is adapted to $\tilde{\sigma}_{k-1}$. With $\tilde{Z}_k \in \tilde{\sigma}_k$, and $\tilde{\sigma}_{k-1} \subset \tilde{\sigma}_k$, $\mathbf{u}_k$ that is determined by $Z_k = \tilde{Z}_k + \bar{Z}_k$ is adapted to $\tilde{\sigma}_k$. ·

Due to the result of Theorem 1, i.e., that the control is adapted to $\tilde{\sigma}_k$, the conditioning on Z_k can be replaced by $\tilde{Z}_k$. With these substitutions, the optimization or maximization step of the model predictive control problem is restated as

$$\begin{aligned} J^*_{\tilde{Z}_k} &= \max_{U_k^{p-1}} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \middle| Z_k\right] \\ &= \max_{U_k^{p-1}} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \middle| \tilde{Z}_k\right] \\ &\text{t:. } \max_{U_k^{p-1}} J_{\tilde{Z}_k} \text{ t:. } J^*_{\tilde{Z}_k}. \end{aligned} \tag{19}$$

In the model predictive control operation mode, although the optimal control sequence is determined over the prediction interval from k to p, only the current control input $\mathbf{u}_k$ at time step k is applied to the system. Then, at subsequent time steps, the performance index in (16) is maximized again to compute a new optimal control sequence, the first element of which is applied to the system.

The conditional pdf needed to evaluate (19) can be determined using the results in Idan and Speyer (2010). Next, we briefly review those results and show how they are used to express $J_{\tilde{Z}_k}$ analytically.

5.1. Propagation of the conditional pdf

In Idan and Speyer (2010) it was shown that the conditional pdf discussed above can be determined analytically in a *closed form*. It was obtained by expressing the pdf as a sum of terms, each of which is a rational function of the random variable. The number of terms in this sum grows with time and the parameters of these terms are updated during the time propagation and measurement updated steps. The *posteriori* form of the conditional pdf given the measurement history up to the current time is presented in (A.12) of Appendix A and presented here for convenience as

$$f_{\tilde{\mathbf{x}}_k|\tilde{\sigma}_k}\left(\tilde{x}_k \middle| \tilde{Z}_k\right) = \sum_{i=1}^{k+1} \frac{a_i(k|k)\tilde{x}_k + b_i(k|k)}{\left(\tilde{x}_k - \lambda_i(k|k)\right)^2 + \omega_i^2(k|k)}, \tag{20}$$

where the conditional pdf is explicitly conditioned on the σ-algebra $\tilde{\sigma}_k$ generated by $\tilde{Z}_k$, the stochastic part (17a) of the decomposed system in (17). The update and propagation of the conditional density, expressed as the propagation equations for $a_i(k|k)$, $b_i(k|k)$, $\lambda_i(k|k)$ and $\omega_i(k|k)$, as well as their boundary conditions, are given in Appendix A. Note that the measurement enters the conditional pdf in a nonlinear way through the $\lambda_{k+2}(k|k) = \tilde{\mathbf{z}}_k/H$ term used in the measurement update stage. In addition, due to the state decomposition of (17), the conditional pdf in (20) is independent of the control variable.

5.2. Construction of performance index

The conditional pdf and its propagation relations reviewed in Appendix A can be used to derive an analytical expression for the conditional performance index J in (19) as a function of the control sequence U_k^{p-1}. First, (6) is used in (19) to yield

$$\begin{aligned} J^*_{\tilde{Z}_k} &= \max_{U_k^{p-1}} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \middle| \tilde{Z}_k\right] \\ &= \max_{U_k^{p-1}} E\left[\prod_{i=k}^{p-1} M_x(\mathbf{x}_{i+1}) M_u(\mathbf{u}_i) \middle| \tilde{Z}_k\right] \\ &= \max_{U_k^{p-1}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=k}^{p-1} M_x(\mathbf{x}_{i+1}) M_u(\mathbf{u}_i) \\ &\quad \times f_{\tilde{\mathbf{x}}_p \cdots \tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_p, \ldots, \tilde{x}_{k+1} \middle| \tilde{Z}_k\right) d\tilde{x}_p, \ldots, d\tilde{x}_{k+1}. \end{aligned} \tag{21}$$

Here, $f_{\tilde{\mathbf{x}}_p \cdots \tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_p, \ldots, \tilde{x}_{k+1} \middle| \tilde{Z}_k\right)$ is the conditional joint density, determined from the one-step projected conditional pdf of (A.17) and the state transition pdf's that are influenced by the process noise, and is determined by

$$\begin{aligned} &f_{\tilde{\mathbf{x}}_p \cdots \tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_p, \ldots, \tilde{x}_{k+1} \middle| \tilde{Z}_k\right) \\ &= f_{\tilde{\mathbf{x}}_p|\tilde{\mathbf{x}}_{p-1} \cdots \tilde{\mathbf{x}}_{k+1}\tilde{\sigma}_k}\left(\tilde{x}_p \middle| \tilde{x}_{p-1}, \ldots, \tilde{x}_{k+1}, \tilde{Z}_k\right) \\ &\quad \times f_{\tilde{\mathbf{x}}_{p-1}|\tilde{\mathbf{x}}_{p-2} \cdots \tilde{\mathbf{x}}_{k+1}\tilde{\sigma}_k}\left(\tilde{x}_{p-1} \middle| \tilde{x}_{p-2}, \ldots, \tilde{x}_{k+1}, \tilde{Z}_k\right) \cdots \\ &\quad \times f_{\tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_{k+1} \middle| \tilde{Z}_k\right) \\ &= f_{\tilde{\mathbf{x}}_p|\tilde{\mathbf{x}}_{p-1}}\left(\tilde{x}_p \middle| \tilde{x}_{p-1}\right) f_{\tilde{\mathbf{x}}_{p-1}|\tilde{\mathbf{x}}_{p-2}}\left(\tilde{x}_{p-1} \middle| \tilde{x}_{p-2}\right) \cdots \\ &\quad \times f_{\tilde{\mathbf{x}}_{k+2}|\tilde{\mathbf{x}}_{k+1}}\left(\tilde{x}_{k+2} \middle| \tilde{x}_{k+1}\right) f_{\tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_{k+1} \middle| \tilde{Z}_k\right), \end{aligned} \tag{22}$$

while from (2) and using the Markov property of (17a), the transition probability is

$$f_{\tilde{\mathbf{x}}_{j+1}|\tilde{\mathbf{x}}_j}\left(\tilde{x}_{j+1} \middle| \tilde{x}_j\right) = \frac{\beta/\pi}{(\tilde{x}_{j+1} - \Phi\tilde{x}_j)^2 + \beta^2} \tag{23}$$

for $j = k, \ldots, p-1$. Substituting (22) into (21) we obtain

$$\begin{aligned} J^*_{\tilde{Z}_k} = \max_{U_k^{p-1} \in F} &\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} M_x(\mathbf{x}_p) M_u(\mathbf{u}_{p-1}) f_{\tilde{\mathbf{x}}_p|\tilde{\mathbf{x}}_{p-1}}\left(\tilde{x}_p \middle| \tilde{x}_{p-1}\right) \\ &\times d\tilde{x}_p \cdots M_x(\mathbf{x}_{k+1}) M_u(\mathbf{u}_k) f_{\tilde{\mathbf{x}}_{k+1}|\tilde{\sigma}_k}\left(\tilde{x}_{k+1} \middle| \tilde{Z}_k\right) d\tilde{x}_{k+1}. \end{aligned} \tag{24}$$

Due to the rational polynomial structure of both the associated pdf's presented earlier and of the objective function $\psi(\cdot, \cdot)$ in (6), the above integrals can be carried out analytically. However, for presentation simplicity, we will derive a closed form result when

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

the state only at the prediction horizon $p = k+m$ is weighted in the objective function of (6). This is obtained by letting $\eta_{i+1} \to \infty$, $i = k, \ldots, p-2$, which yields $M_x(\mathbf{x}_{i+1}) = 1$, $i = k, \ldots, p-2$. In this case the conditional performance index to maximize becomes

$$J^*_{\tilde{z}_k} = \max_{U_k^{p-1}} \prod_{i=k}^{p-1} M_u(\mathbf{u}_i) \int_{-\infty}^{\infty} M_x(\mathbf{x}_p) f_{\tilde{x}_p|\tilde{\sigma}_k}\left(\tilde{x}_p|\tilde{Z}_k\right) d\tilde{x}_p. \tag{25}$$

The conditional pdf $f_{\tilde{x}_p|\tilde{\sigma}_k}\left(\tilde{x}_p|\tilde{Z}_k\right)$ is obtained by time-propagating $f_{\tilde{x}_k|\tilde{\sigma}_k}\left(\tilde{x}_k|\tilde{Z}_k\right)$ of (20) m times using (A.17)–(A.21) to yield

$$f_{\tilde{x}_p|\tilde{\sigma}_k}\left(\tilde{x}_p|\tilde{Z}_k\right) = \prod_{i=1}^{k+2} \frac{a_i(p|k)\tilde{x}_p + b_i(p|k)}{\tilde{x}_p - \lambda_i(p|k)^2 + \omega_i^2(p|k)}. \tag{26}$$

The parameters $a_i(p|k)$, $b_i(p|k)$, $\lambda_i(p|k)$ and $\omega_i(p|k)$ are expressed explicitly using (A.18)–(A.21). The integral in (25) is evaluated analytically as

$$\begin{aligned} g(\bar{\mathbf{x}}_p) &= \int_{-\infty}^{\infty} M_x(\mathbf{x}_p) f_{\tilde{x}_p|\tilde{\sigma}_k}\left(\tilde{x}_p|\tilde{Z}_k\right) d\tilde{x}_p \\ &= \int_{-\infty}^{\infty} \frac{\eta_p^2}{(\tilde{x}_p + \bar{\mathbf{x}}_p)^2 + \eta_p^2} \prod_{i=1}^{k+2} \frac{a_i(p|k)\tilde{x}_p + b_i(p|k)}{\tilde{x}_p - \lambda_i(p|k)^2 + \omega_i^2(p|k)} d\tilde{x}_p \\ &= \prod_{i=1}^{k+2} \int_{-\infty}^{\infty} \frac{\eta_p^2}{(\tilde{x}_p + \bar{\mathbf{x}}_p)^2 + \eta_p^2} \frac{a_i(p|k)\tilde{x}_p + b_i(p|k)}{\tilde{x}_p - \lambda_i(p|k)^2 + \omega_i^2(p|k)} d\tilde{x}_p \\ &= \pi\eta_p^{k+2} \prod_{i=1}^{\omega_i} \frac{1}{(\bar{\mathbf{x}}_p + \lambda_i)^2 + (\eta_p + \omega_i)^2} \frac{a_i(\bar{\mathbf{x}}_p + \lambda_i)\eta_p + (b_i - a_i\bar{\mathbf{x}}_p)(\eta_p + \omega_i)}{\omega_i}, \end{aligned} \tag{27}$$

where in the last expression the functional dependence of the parameters a_i, b_i, λ_i and ω_i on the time index $(p|k)$ is removed for brevity. Using this result in (25) yields the following maximization problem

$$J^*_{\tilde{z}_k} = \max_{U_k^{p-1}} \prod_{j=k}^{p-1} \frac{\zeta_j^2}{\mathbf{u}_j^2 + \zeta_j^2} g(\bar{\mathbf{x}}^*_p). \tag{28}$$

The dependence of $J_{\tilde{Z}_k}$ on the elements of U_k^{p-1} is given explicitly in the first term of (28) and implicitly in the second term via $\bar{\mathbf{x}}_p$, which is obtained by propagating (17b) m times using the control sequence U_k^{p-1}; hence, $\bar{\mathbf{x}}_p$ is a function of an open loop control policy from time k to time $p-1$, and $\bar{\mathbf{x}}^*_p$ is the maximizing value. Note that, as with the *a priori* conditional pdf in (20), the number of terms that determine the performance index in (28) grows with k.

The conditional performance index in (28) is an extremely complex function of the m elements in U_k^{p-1}. Consequently, there is no analytical solution to the associated maximization problem. Hence, a numerical procedure is proposed next to determine the optimal control signal at each time step k. Since the numerics use realizations where random variables take on specific values, regular fonts are used hereon for the state and control.

5.3. Homotopy optimization for maximizing $J_{\tilde{Z}_k}$

To solve the maximization problem in (28) using the homotopy method, we first address the problem when the control variable is removed from the performance index, attained by letting $\zeta_j \to \infty$, $j = k, \ldots, p-1$. In this case, $J_{\tilde{Z}_k}$ becomes a sum of rational functions of $\bar{x}_p$. Since $\bar{x}_p$ is a scalar, a one dimensional search can be used to determine the global optimum, $\bar{x}^*_p$. Since there are no control penalties, this global optimum is focused on keeping the predicted state small for regulation. Moreover, due to the deterministic state propagation in (17b), for $m > 1$ the control sequence that yields this particular $\bar{x}^*_p$ is not unique (clearly it is unique when $m = 1$.) One such sequence, which assumes that u_j, $j = k, \ldots, p-1$ is constant, is given by

$$u_j = \frac{\bar{x}^*_p - \Phi^m \bar{x}_k}{\sum_{i=0}^{m-1} \Phi^i}, \quad j = k, \ldots, p-1. \tag{29}$$

Next, we use homotopy to incorporate control penalties back into the performance index. For that, the weights ζ_j, $j = k, \ldots, p-1$ are set to very high values, thus imposing a small change in the value of the performance index. Starting with an initial guess for u_j, $j = k, \ldots, p-1$ as in (29), the maximization of $J_{\tilde{Z}_k}$ is carried out numerically using the accelerated gradient method (Fletcher & Powell, 1963; Myers, 1968) as a refinement step. The process is repeated for decreasing ζ_js until reaching their design values.

The gradient needed for the accelerated gradient method in the refinement step in the homotopy optimization method can be computed analytically by applying the chain rule to (28). The partial of the performance index $J_{\tilde{Z}_k}$ with respect to u_ℓ is given by

$$\begin{aligned} \frac{\partial J_{\tilde{Z}_k}}{\partial u_\ell} &= \prod_{j=k, j\neq\ell}^{p-1} \frac{\zeta_j^2}{u_j^2 + \zeta_j^2} \frac{\zeta_\ell^2}{u_\ell^2 + \zeta_\ell^2} \\ &\times \Phi^{p-\ell} \frac{\partial g(\bar{x}_p)}{\partial \bar{x}_p} - \frac{2u_\ell}{u_\ell^2 + \zeta_\ell^2} g(\bar{x}_p) \end{aligned} \tag{30}$$

for $\ell = k, \ldots, p-1$, where $g(\bar{x}_p)$ is given in (27) and its gradient is

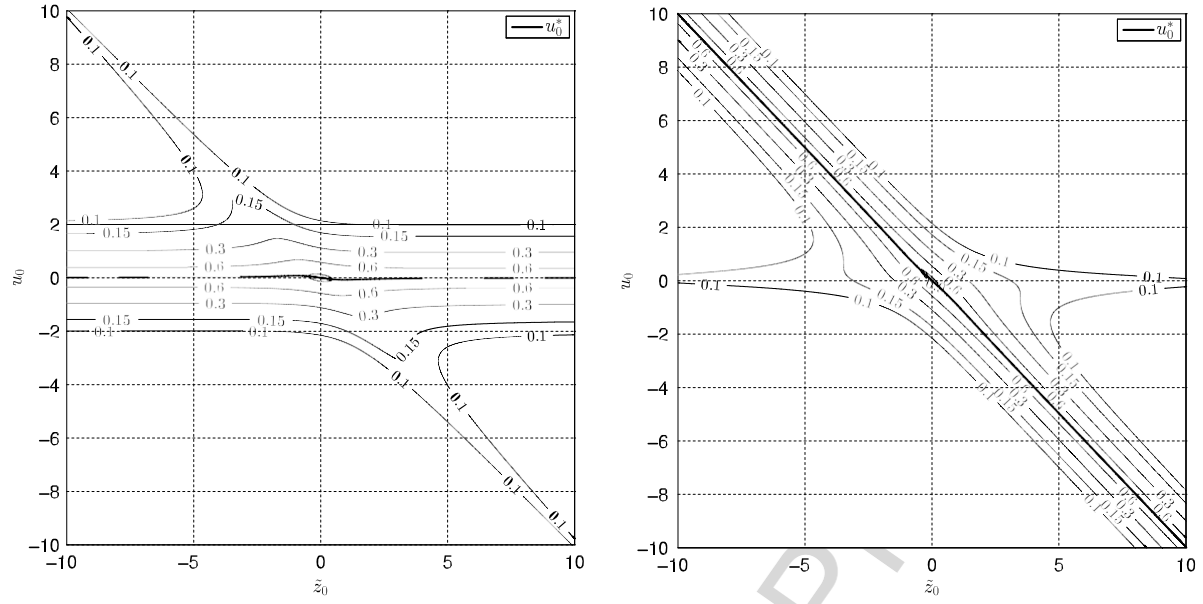
$$\begin{aligned} \frac{\partial g(\bar{x}_p)}{\partial \bar{x}_p} &= \pi\eta_p^{k+2} \prod_{i=1}^{a_i} \frac{\bar{x}_p^2 - \lambda_i^2 + \eta_p + \omega_i}{(\bar{x}_p + \lambda_i)^2 + (\eta_p + \omega_i)^2} \\ &- 2\bar{x}_p + \lambda_i \frac{a_i\lambda_i\eta_p + b_i\eta_p + \omega_i}{\omega_i (\bar{x}_p + \lambda_i)^2 + (\eta_p + \omega_i)^2}. \end{aligned} \tag{31}$$

Although an analytic form of the second derivative of $g(\bar{x}_p)$ can be determined, it involves a sum of terms that increases the computation time. Thereby, we use an accelerated gradient method that iteratively estimates the Hessian matrix of the performance index from an initial guess, and by performing a sequence of one-dimensional searches. Each such search maximizes the function along a specific search direction, determined by the gradient and the current estimate of the Hessian. The search result is used to update the estimate of the Hessian, which is then used to determine a new search direction orthogonal to the previous one. For a quadratic performance index this second order method converges to the optimum in n steps, where n is the number of variables being optimized over (Fletcher & Powell, 1963; Myers, 1968). Although the Cauchy performance index is not quadratic, it is approximately quadratic locally around the maximum. Therefore, once a point sufficiently near the maximum is found, the accelerated gradient method will find this maximum easily. To find this start point for the accelerated gradient method, the homotopy optimization method uses a grid search to find the optimal $\bar{x}^*_p$ and then finds a predictive control sequence to get from $\bar{x}_k$ to $\bar{x}^*_p$. From that point, we do a simple gradient search to get to the quadratic neighborhood before using the accelerated gradient method to find the optimum.

6. Numerical examples

To obtain insight into the properties of the Cauchy stochastic controller, the one-step, one-measurement example is first

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005



(a) $\alpha = 0.1,\ \beta = 0.02,\ \gamma = 0.5,\ \eta_1 = 0.7,\ \zeta_0 = \infty$. (b) $\alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,\ \eta_1 = 0.7,\ \zeta_0 = \infty$.

Fig. 1. Contour plots of the performance index for the Cauchy one-step controller.

analyzed in Section 6.1. Then, multi-step numerical simulation results are given in Section 6.2, which illustrate the Cauchy controllers behavior in the presence of dominant Cauchy measurement noise and then dominant Cauchy process noise. The performance of the Cauchy controller is compared with a somewhat standard scheme, here the model predictive linear-exponential-Gaussian (LEG) controller. Although dynamic programming provides a closed-form solution to the LEG problem (Speyer & Chung, 2008, Chapter 10), the m-step model predictive controller is a bit different and its solution process is much simpler to obtain. The derivation of the LEG m-step controller, used here for comparison, is given in Appendix B. Also, in order to compare the Cauchy controller with the LEG controller, a least-squares fit between these two classes of pdf's, and thus the objective functions, is given in Appendix B.1 and used in the subsequent examples.

6.1. One-step, one-measurement example

To explore the characteristics of the proposed Cauchy controller, in this section we evaluate numerically its simplest form. Specifically, we examine the value of the optimal control signal at $k = 0$, i.e., $u_0^*(\tilde{z}_0)$, as a function of the first measurement $\tilde{z}_0$, that varies due to the measurement noise v_0 (Idan et al., 2010), while considering a one step horizon, i.e., $m = 1$. The parameters for the system and Cauchy signals are first chosen as $\Phi = 1$, $H = 1$, $\alpha = 0.1$, $\beta = 0.02$, $\gamma = 0.5$, and $\bar{x}_0 = 0$. This example represents the case where the uncertainty in the initial condition is smaller than the measurement noise, i.e., $\alpha < \gamma$. To explore the effect of higher uncertainty in the initial conditions, also the case for which $\alpha > \gamma$ is considered by choosing the values $\alpha = 0.5$ and $\gamma = 0.1$.

Initially, no penalty is introduced on the control signal in (28), i.e., $\zeta_0 \to \infty$, while the state at $k = 1$ is weighted with $\eta_1 = 0.7$. Substituting the system parameters into (28), the performance index becomes

$$J^*_{\tilde{z}_0} = \frac{0.1148(4.1667\tilde{z}_0^2 - 1.0163u_0\tilde{z}_0 + 1)}{(\tilde{z}_0^2 + 0.16)(u_0^2 + 0.6724)} + \frac{0.03416(7.5820\tilde{z}_0^2 + 3.4153u_0\tilde{z}_0 - 1)}{(\tilde{z}_0^2 + 0.16)\left[(u_0 + \tilde{z}_0)^2 + 1.22\right]}. \quad (32)$$

The optimal controller can be obtained by maximizing (32) with respect to u_0. The necessary optimality condition, $\partial J^*_{\tilde{z}_0}/\partial u_0 = 0$, reduces to finding the roots of the fifth-order polynomial

$$l_5 u_0^5 + l_4 u_0^4 + l_3 u_0^3 + l_2 u_0^2 + l_1 u_0 + l_0 = 0, \quad (33)$$

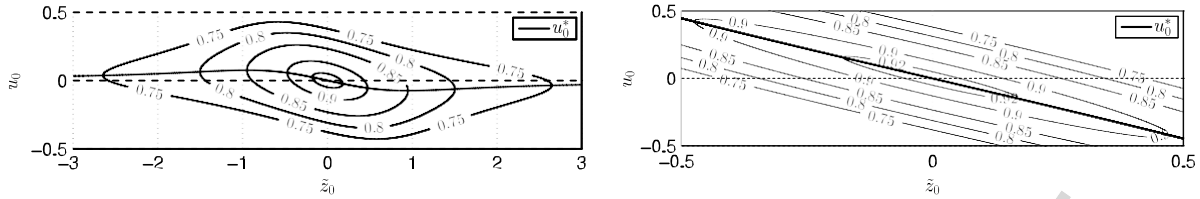
where

$$\begin{aligned} l_5 &= 1, \quad l_4 = 3.5\tilde{z}_0, \quad l_3 = (5.2315\tilde{z}_0^2 + 3.6681), \\ l_2 &= (3.6806\tilde{z}_0^3 + 6.6305)\tilde{z}_0, \\ l_1 &= (0.9491\tilde{z}_0^4 + 3.2124\tilde{z}_0^2 + 2.9623), \\ l_0 &= (0.07782\tilde{z}_0^3 + 0.3992)\tilde{z}_0. \end{aligned} \quad (34)$$

This fifth order polynomial always has at least one real root. If three roots are real, then there are two local maximum values and the larger of the two gives the optimal control. A similar expression, with different numerical values, can be attained also for the $\alpha > \gamma$ case.

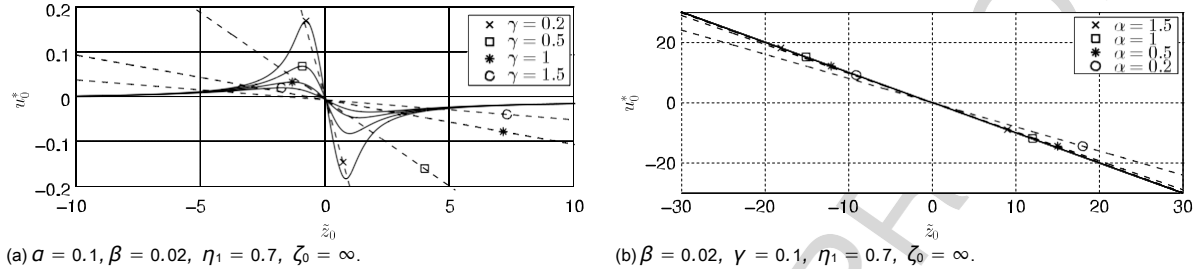
Contour plots of the performance index for the two case discussed above are shown in Fig. 1, with expanded views given in Fig. 2. In both cases it is observed that the performance index has two ridges, with one, marked by the dashed line, being dominant. This indicates that the solution of (33) had two maxima, with the higher one being the global maximum. These contours demonstrate a clear difference in the optimal control action for the two cases examined. For the first case of $\alpha < \gamma$, shown in Figs. 1(a) and 2(a), the optimal control action u^*_0 is a nonlinear function of $\tilde{z}_0$ and is almost aligned with the $u_0 \approx 0$ line. In the center is an ellipse shaped contour whose major axis is oriented between the two ridges. This demonstrates that when the measurement uncertainty is dominant, the optimal control action is minimal. The optimal control strategy changes drastically for the $\alpha > \gamma$ case depicted in Figs. 1(b) and 2(b). In this case, the global optimum generates an approximately linear relation between u^*_0 and $\tilde{z}_0$, indicating that a significant control action is adopted when the measurement uncertainty is decreased.

This difference in behavior cannot be deduced by examining only the conditional variance and conditional mean of the estimation error at $k = 0$, which are given in Idan and Speyer



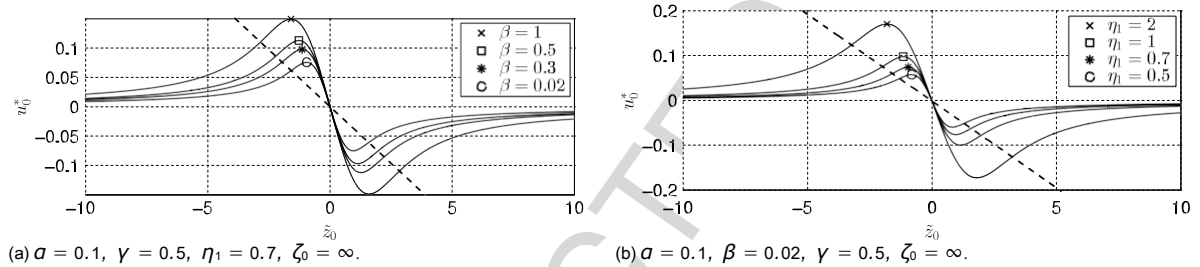
(a) $\alpha = 0.1,\ \beta = 0.02,\ \gamma = 0.5,\ \eta_1 = 0.7,\ \zeta_0 = \infty$. (b) $\alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,\ \eta_1 = 0.7,\ \zeta_0 = \infty$.

Fig. 2. Zoomed in contour plots of the performance index for the Cauchy one-step controller.



(a) $\alpha = 0.1, \beta = 0.02,\ \eta_1 = 0.7,\ \zeta_0 = \infty$. (b) $\beta = 0.02,\ \gamma = 0.1,\ \eta_1 = 0.7,\ \zeta_0 = \infty$.

Fig. 3. Cauchy (solid line) and Gaussian (dashed line) one-step controller with parameters variations in α and γ.



(a) $\alpha = 0.1,\ \gamma = 0.5,\ \eta_1 = 0.7,\ \zeta_0 = \infty$. (b) $\alpha = 0.1,\ \beta = 0.02,\ \gamma = 0.5,\ \zeta_0 = \infty$.

Fig. 4. Cauchy (solid line) and Gaussian (dashed line) one-step controller with parameters variations in β and η_1.

(2008) as

$$E\left[(x_0 - \hat{x}_0)^2 | \tilde{z}_0\right] = \alpha\gamma\left[\frac{\tilde{z}_0^2}{(\alpha+\gamma)^2} + 1\right], \tag{35}$$

where $\hat{x}_0 = \alpha\gamma\tilde{z}_0/(\alpha+\gamma)$. This conditional variance represents the uncertainty in the state estimation and grows with $\tilde{z}_0^2$. Eq. (35) shows that the values of the conditional variance and the conditional mean do not change when interchanging the values of γ and α. Hence, the change in the control strategy is correctly deduced by the optimal Cauchy control scheme. Specifically, it correctly reduces the control effort when $\alpha < \gamma$, i.e., when a measurement noise impulse is more likely than a large impulse in the initial condition. Alternatively, a nearly linear control action is produced when $\alpha > \gamma$, i.e., the controller effectively forces the state back toward the origin when the measurement uncertainty is relatively small.

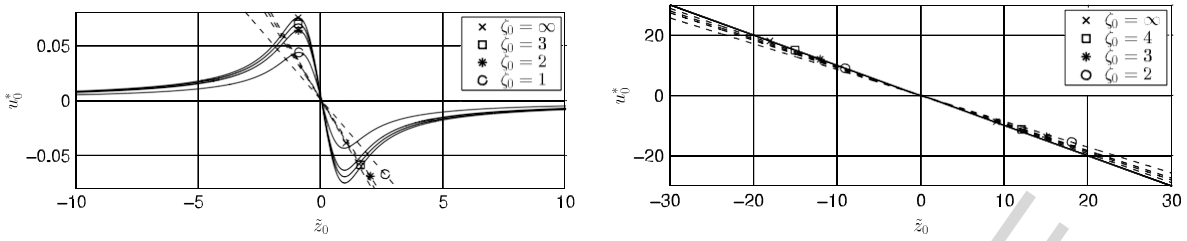
To obtain further insight, the performance of the Cauchy controller is compared to the LEG controller, presented in Appendix B. First, in Fig. 3(a), the optimal Cauchy and LEG control signals are plotted versus the measurement $\tilde{z}_0$ for the $\alpha < \gamma$ case with different values of γ. In this case, for the LEG controller, the optimal control is given by $u^*_0 = -\hat{x}_0$ and therefore, it is linear in $\tilde{z}_0$. Fig. 3(a) shows that the Cauchy controller is symmetric about $\tilde{z}_0 = 0$. Furthermore, the Cauchy controller approaches zero when $|\tilde{z}_0|$ becomes large, while the Gaussian controller remains linear with respect to the measurement $\tilde{z}_0$. This is a significant difference in behavior between the Cauchy and Gaussian optimal controllers that can be deduced analytically from (33). If $u^*(\tilde{z}_0)$ is finite, the dominant term in (33) as $|\tilde{z}_0| \to \infty$ is $l_1 u^*(\infty)$, or $\lim_{|\tilde{z}_0|\to\infty} u^*(\tilde{z}_0) \to 0$.

Fig. 3(b) examines the case where $\gamma < \alpha$ for different values of α. In this case also the Cauchy controller is nearly linear. Interestingly, the slope of the Cauchy controller, which is hardly effected by the change α, is nearly identical to that of the LEG controller.

The effect of parametric changes in β and η_1 for the $\gamma > \alpha$ case is explored in Fig. 4. In this case, the LEG controller is a single line, since for $\zeta_0 = \infty$, i.e., with no control weighting, the LEG controller is independent of the parameters β and η_1 (see Appendix B). Exploring this parametric change in the $\gamma < \alpha$ reveals plots similar to Fig. 3(b), not shown here for brevity.

Now we consider the cases with control weighting. Reducing ζ_0 from ∞ to 5 has a negligible effect on both the Cauchy and the LEG controllers. Further reduction in ζ_0, i.e., increase in the control weighting, affects the control strategies as depicted in Fig. 5. The nonlinear behavior of the Cauchy controller is retained for the $\gamma > \alpha$ case depicted in Fig. 5(a), while an approximately linear Cauchy controller is shown in Fig. 5(b) for the $\gamma < \alpha$.

For $\gamma > \alpha$, in the region where $\tilde{z}_0$ is relatively small, the Cauchy controller is approximately linear (see Fig. 5(a)). However, the Cauchy controller in Fig. 5(a) for large $\tilde{z}_0$ goes toward zero. This is in sharp contrast with the LEG controller, which remains linear in the measurement. Therefore, the problem of handling outliers, which occur for the Cauchy pdf, appears to be resolved by the Cauchy controller explicitly, and not in the filtering or estimation stage as has been done traditionally. Note that the controller design process explicitly uses the parameters γ and α and hence their relative size $\gamma > \alpha$, i.e., it anticipates more impulsive measurement uncertainty than process uncertainty. If $\gamma < \alpha$, the Cauchy controller behaves approximately like the LEG



(a) $\alpha = 0.1,\ \beta = 0.02,\ \gamma = 0.5,\ \eta_1 = 0.7.$ (b) $\alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,\ \eta_1 = 0.7.$

Fig. 5. Cauchy (solid line) and Gaussian (dashed line) one-step controller with parameters variations in ζ_0 for $\gamma > \alpha$ left and $\gamma < \alpha$ right.

linear controller in Fig. 5(b), i.e., it expects more impulsive process uncertainty than measurement uncertainty. For both the Cauchy and Gaussian changes in ζ_0 make only small changes in the gain.

6.2. Multi-step numerical examples

The dynamic characteristics of the Cauchy optimal controller, obtained by maximizing the performance index in (28), are explored through several multi-step numerical examples. The Cauchy optimal control results are compared against the least-squares equivalent LEG controller, obtained from (B.12) and the Kalman filter from (B.10). The two examples that are discussed in this section are a stable system with a horizon length of $m = 2$ and an unstable system with a horizon length of $m = 5$. The parameter values of β and γ are interchanged to see how the controller performance changes when it is designed for a large measurement noise impulse in contrast to when it is designed for a large process noise impulse. All simulations in this section use either noise parameters $\beta = 0.1, \gamma = 0.02$ or $\beta = 0.02, \gamma = 0.1$ and initial condition: $\alpha = 0.5$ and $\bar{x}_0 = 0$. They also all use the same system parameter of $H = 1$, and either $\Phi = 0.95$ or $\Phi = 1.05$. Substituting these parameters into (28), the performance index is maximized numerically with respect to the control at each time using the homotopy optimization method.

The performance of the Cauchy and LEG model predictive controllers are compared for $m = 2$ and using $\eta_p = 0.7$ as the parameter for the penalty on the state at x_{k+2} and $\zeta_i = 8, i = k, k+1$ as the parameter for the control penalty. First, for $\gamma = 0.1$ and $\beta = 0.02$, when the noises are small, the Cauchy and the LEG controllers exhibit similar performance. However, they behave rather differently when a large measurement pulse occurs. A significant measurement noise pulse causes a large measurement. A measurement noise pulse does not represent a state deviation and thus, for proper regulation, the controller should ignore that measurement. The Cauchy predictive controller, designed for $\gamma > \beta$, is able to make this distinction, whereas the LEG predictive controller reacts linearly to all the pulses and does not differentiate as shown in Fig. 6(a), at around time step $k = 51$. The Cauchy controller ignores this large measurement deviation, applying almost zero control, whereas the LEG controller applies a very large control input that causes the state to deviate away from zero. An additional control effort is then required to correct this deviation. In this way the Cauchy controller manages to avoid unnecessary actuation and thus maintains the system performance.

Performance differences are also observed when encountering a large process noise signal. In Fig. 6(a), at time steps $k = 1$ and $k = 13$ process noise pulses occur, and although both controllers react to them and are able to overcome this deviation, the Cauchy controller does so much quicker than the LEG by applying a much larger control effort. The Cauchy controller applies a larger control because its gain for small measurement values are higher than that of the LEG.

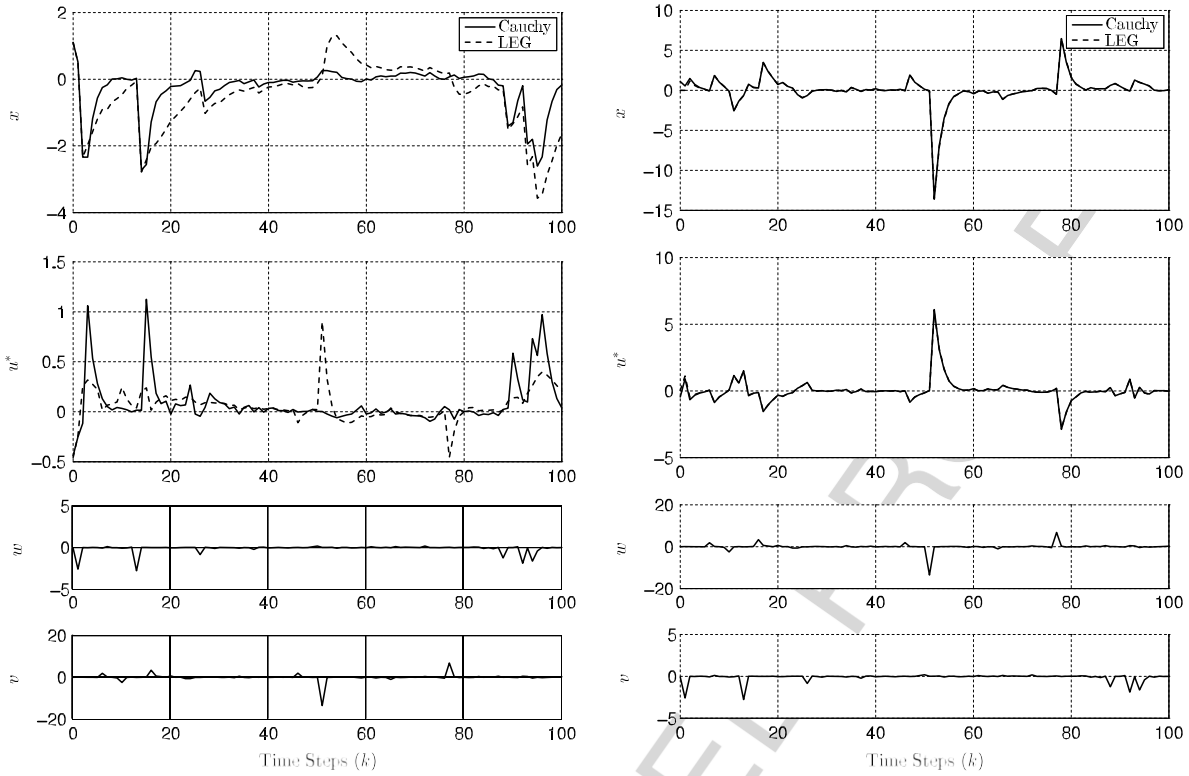
When the Cauchy predictive controller is designed for $\beta > \gamma$, the behavior of both the Cauchy and LEG controllers perform similarly as shown in Fig. 6(b), demonstrating the same linear behavior seen in Figs. 3(b) and 5(b). One effect of having $\beta > \gamma$ is that the state trajectories for the Cauchy and LEG controllers appear equal in Figs. 6(b) and 7(b). Even though the controllers process the measurements differently, due to the dominance of the process noise the controls and state appear similar.

Using the same noise sequence, similar results are obtained for an unstable system with $m = 5$, while using $\eta_p = 0.7$ as the parameter for the penalty on the state at x_{k+5} and $\zeta_i = 8, i = k, \ldots, k+4$ as the parameter for the control penalty. These results are depicted in Fig. 7. For the $\gamma > \beta$ shown in Fig. 7(a), the measurement noise impulses are more probable than process noise impulses. Hence, when a large measurement impulse occurs at $k = 51$, the Cauchy controller ignores it, applying almost zero control, whereas the LEG controller applies a very large control input that causes the state to deviate away from zero. For the $\gamma < \beta$ case shown in Fig. 7(b), both the Cauchy and the LEG controllers regulate in a similar fashion.

Finally, the Cauchy controller is examined in simulation with Gaussian noises, for which the LEG design is optimal. The design parameters used for the Cauchy controller were obtained by using the least-square fit relations between the Cauchy and Gaussian pdf's presented in Appendix B. Both a stable system with $m = 2$ and an unstable system with $m = 5$ presented earlier were used in these simulations. The results are presented in Fig. 8, showing also the numerical values used in the Cauchy controller design. The response of the Cauchy controller is very similar to that of the optimal (in this case) LEG controller for both cases. This clearly demonstrates the robustness of the Cauchy controller, which performs nearly optimally even in the Gaussian noise simulation for which it was not designed.

7. Conclusions

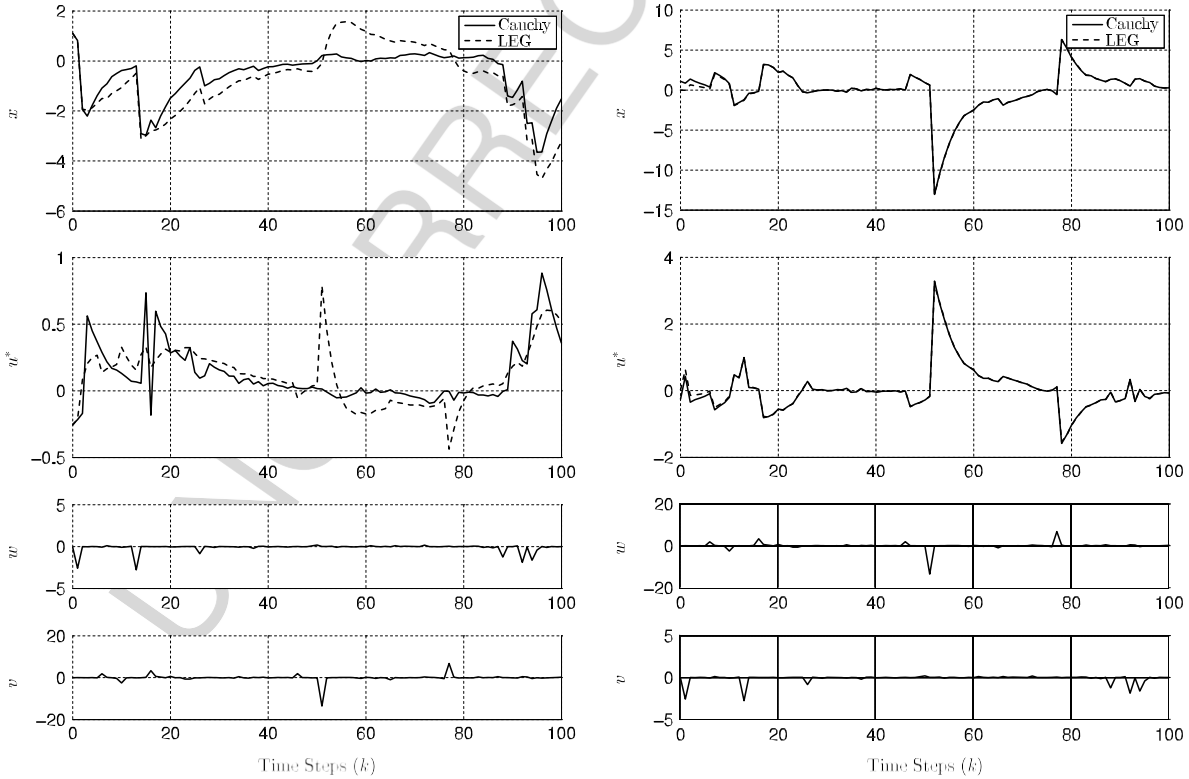
A new control design paradigm is proposed for optimal control problems for scalar linear dynamic system driven by Cauchy distributed signals. The scalar dynamic programming solution is intractable. Since the model predictive controller is known to have similar, though suboptimal, performance, an m-step optimal model predictive controller was derived. The conditional expectation of the new objective function with respect to the exact Cauchy conditional pdf was determined in closed form. The Cauchy controller results were compared against an equivalent LEG m-step optimal model predictive controller. We showed that both controllers behave in a similar way when faced with non-impulsive Gaussian noises that generate small measurement noise. However, a dramatic difference was observed between the two controllers when faced with Cauchy measurement noise sequences with large pulses. The Cauchy optimal controller ignores measurement noise pulses and reacts strongly to process noise pulses, effectively differentiating between process and measurement noise and thereby producing a scheme for handling measurement outliers. This



(a) $\alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,\ \Phi = 0.95,\ \eta_p = 0.7,$ $\zeta_i = 8,\ m = 2.$

(b) $\alpha = 0.5,\ \beta = 0.1,\ \gamma = 0.02,\ \Phi = 0.95,\ \eta_p = 0.7,$ $\zeta_i = 8,\ m = 2.$

Fig. 6. 2-step Cauchy and Gaussian controllers with β and γ parameters interchanged for stable system in Cauchy simulation.

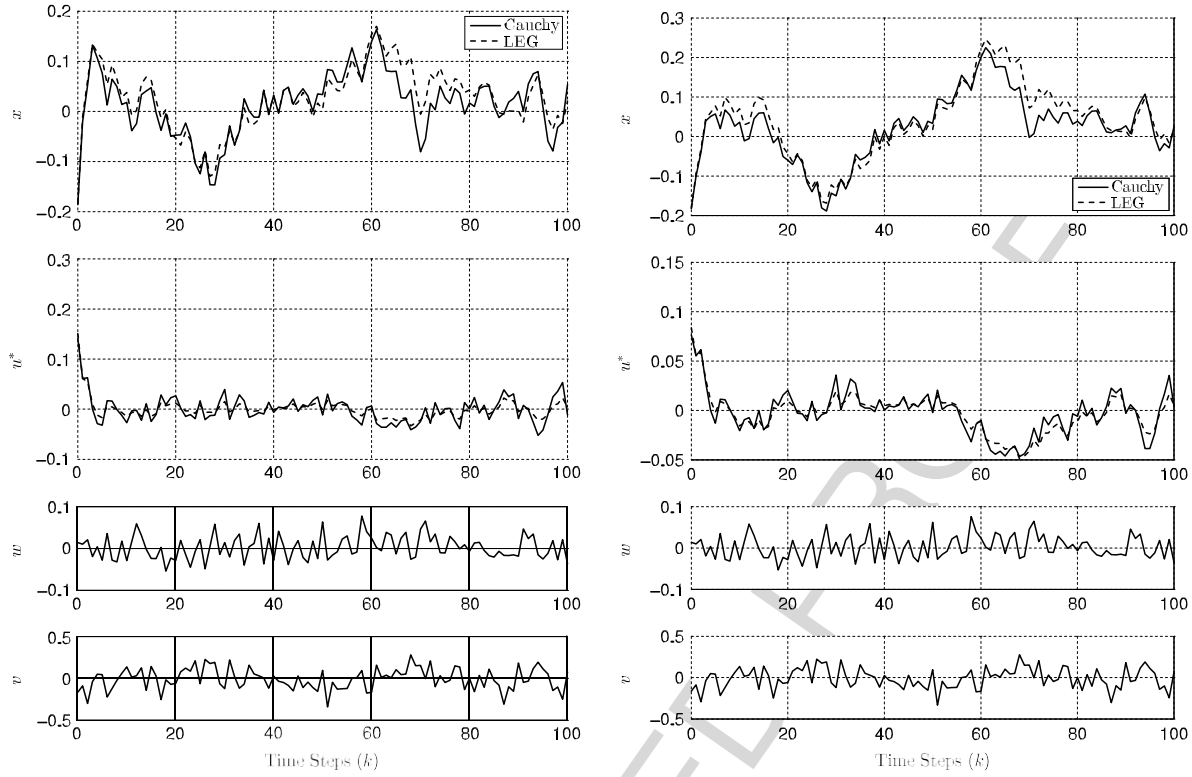


(a) $\alpha = 0.5,\ \beta = 0.02,\ \gamma = 0.1,\ \Phi = 1.05,\ \eta_p = 0.7,$ $\zeta_i = 8,\ m = 5.$

(b) $\alpha = 0.5,\ \beta = 0.1,\ \gamma = 0.02,\ \Phi = 1.05,\ \eta_p = 0.7,$ $\zeta_i = 8,\ m = 5.$

Fig. 7. 5-step Cauchy and Gaussian controllers with β and γ parameters interchanged for unstable system in Cauchy simulation.

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005



(a) $\alpha = 0.5$, $\beta = 0.02$, $\gamma = 0.1$, $\Phi = 0.95$, $\eta_p = 0.7$, $\zeta_i = 8$, $m = 2$.

(b) $\alpha = 0.5$, $\beta = 0.02$, $\gamma = 0.1$, $\Phi = 1.05$, $\eta_p = 0.7$, $\zeta_i = 8$, $m = 5$.

Fig. 8. Cauchy and Gaussian controllers for stable and unstable systems in a Gaussian simulation.

ability of the Cauchy controller to reject outliers depends crucially on the dominance of the measurement noise over the process noise. In contrast, the LEG controller is linearly proportional to the measurements and always reacts to large measurement outliers.

The conditional performance index is a sum that grows at each measurement update. However, as justified in Idan and Speyer (2010), the number of terms in the conditional pdf can be truncated with minimal effect on the performance; in fact, the truncation of old terms enhances numerical performance for unstable systems. This result is important for the feasibility of the Cauchy model predictive controller because it leads to a computationally efficient implementation. Finally, generalization of this scheme is being resolved by first extending the estimator to the multivariable case, where the characteristic function of the conditional pdf must be propagated (Idan & Speyer, 2013). Then, Parseval's Theorem is used to express the conditional performance index as an integral over the spectral variables involving the characteristic function of the conditional pdf (Fernández, Speyer, & Idan, 2013). This formulation also allows a pathway for the generalization of the scalar stochastic controller developed here to the multivariable case.

Appendix A. Algorithm for the propagation of the conditional pdf

This appendix briefly summarizes the Cauchy estimator derived previously in Idan and Speyer (2010). This estimator is central in deriving the Cauchy controller discussed in this study.

The conditional pdf of $\tilde{\mathbf{x}}_k$ given past data $\tilde{Z}_{k-1}$ (i.e., before the $\tilde{\mathbf{z}}_k$ measurement is processed) is shown in Idan and Speyer (2010) to be expressed in a factored form given by

$$f_{\tilde{\mathbf{x}}_k|\tilde{\sigma}_{k-1}}(\tilde{x}_k|\tilde{Z}_{k-1}) = \prod_{i=1}^{k+1} \frac{a_i(k|k-1)\tilde{x}_k + b_i(k|k-1)}{(\tilde{x}_k - \lambda_i(k|k-1))^2 + \omega_i^2(k|k-1)}. \quad \text{(A.1)}$$

At the initial $k = 0$ there is only one term in the above sum, with the initial parameters $a_1(0|-1) = 0$, $b_1(0|-1) = \alpha/\pi$, $\lambda_1(0|-1) = 0$, and $\omega_1(0|-1) = \alpha$. After an additional measurement $\tilde{\mathbf{z}}_k$ is obtained, a measurement updated conditional pdf is determined by

$$f_{\tilde{\mathbf{x}}_k|\tilde{\sigma}_k}(\tilde{x}_k|\tilde{Z}_k) = \frac{f_{\tilde{\mathbf{x}}_k,\tilde{\mathbf{z}}_k|\tilde{\sigma}_{k-1}}(\tilde{x}_k,\tilde{z}_k|\tilde{Z}_{k-1})}{f_{\tilde{\mathbf{z}}_k|\tilde{\sigma}_{k-1}}(\tilde{z}_k|\tilde{Z}_{k-1})}. \quad \text{(A.2)}$$

The density function in the above numerator is computed as

$$f_{\tilde{\mathbf{x}}_k,\tilde{\mathbf{z}}_k|\tilde{\sigma}_{k-1}}(\tilde{x}_k,\tilde{z}_k|\tilde{Z}_{k-1}) = f_{\tilde{\mathbf{x}}_k|\tilde{\sigma}_{k-1}}(\tilde{x}_k|\tilde{Z}_{k-1})\, f_{V_k}(\tilde{z}_k - H\tilde{x}_k)$$
$$= \prod_{i=1}^{k+1} \frac{a_i(k|k-1)\tilde{x}_k + b_i(k|k-1)}{(\tilde{x}_k - \lambda_i(k|k-1))^2 + \omega_i^2(k|k-1)} \times \frac{\gamma/\pi}{(\tilde{z}_k - H\tilde{x}_k)^2 + \gamma^2}. \quad \text{(A.3)}$$

A simple assumption made in Idan and Speyer (2010) guarantees that the complex conjugate roots in (A.3) of the denominator polynomial in $\tilde{x}_k$ are distinct. Therefore, using partial fraction expansions, (A.3) can be written as

$$f_{\tilde{\mathbf{x}}_k,\tilde{\mathbf{z}}_k|\tilde{\sigma}_{k-1}}(\tilde{x}_k,\tilde{z}_k|\tilde{Z}_{k-1}) = \sum_{i=1}^{k+2} \frac{a_i(k|k)\tilde{x}_k + b_i(k|k)}{(\tilde{x}_k - \lambda_i(k|k))^2 + \omega_i^2(k|k)}. \quad \text{(A.4)}$$

This step introduces an additional term in the pdf sum with four parameters. $\lambda_i(k|k) = \lambda_i(k|k-1)$, $\omega_i(k|k) = \omega_i(k|k-1)$, $i = 1, \ldots, k+1$ represent the unchanged roots of the pdf denominator, while $\lambda_{k+2}(k|k) = \tilde{z}_k/H$ and $\omega_{k+2}(k|k) = \gamma/|H|$ represent its new complex measurement dependent roots. $a_i(k|k)$ and $b_i(k|k)$ are computed from the partial fraction expansions as

$$\begin{pmatrix} a_i(k|k) \\ b_i(k|k) \end{pmatrix} = F_i(k) \begin{pmatrix} a_i(k|k-1) \\ b_i(k|k-1) \end{pmatrix} \quad \text{(A.5)}$$

for $i = 1, \ldots, k+1$, where

$$F_i(k) = \frac{1}{\Delta^i(k)} \times \left[\delta_i(k) - \frac{\lambda_i(k|k)}{\omega_i(k|k)}\theta_i(k) \quad -\frac{1}{\omega_i(k|k)}\theta_i(k)\right] \left[\frac{\lambda_i^2(k|k)+\omega_i^2(k|k)}{\omega_i(k|k)}\theta_i(k) \quad \delta_i(k) + \frac{\lambda_i(k|k)}{\omega_i(k|k)}\theta_i(k)\right] \tag{A.6}$$

$$\delta_i(k) = (\lambda_{k+2}(k|k) - \lambda_i(k|k))^2 + \omega_{k+2}^2(k|k) - \omega_i^2(k|k) \tag{A.7a}$$

$$\theta_i(k) = 2\omega_i(k|k)\,(\lambda_i(k|k) - \lambda_{k+2}(k|k)) \tag{A.7b}$$

$$\Delta_i(k) = \frac{\pi\gamma}{\omega_{k+2}^2(k|k)}\left[\delta_i(k) + \theta_i^2(k)\right]. \tag{A.7c}$$

Note that the matrices $F_i(k)$, $i = 1, \ldots, k+1$ are not functions of the numerator parameters $a_i(k|k-1)$ and $b_i(k|k-1)$, generating a linear, time-dependent, stochastic update equation (A.5). The numerator parameters of the new term in the pdf sum are

$$\bar{a}_{k+2}(k|k) = \prod_{i=1}^{k+1} \bar{a}_{i,k+2}(k|k) \tag{A.8}$$

$$\bar{b}_{k+2}(k|k) = \prod_{i=1}^{k+1} \bar{b}_{i,k+2}(k|k), \tag{A.9}$$

where for $i = 1, \ldots, k+1$

$$\begin{bmatrix}\bar{a}_{i,k+2}(k|k)\\ \bar{b}_{i,k+2}(k|k)\end{bmatrix} = F_{i,k+2}(k)\begin{bmatrix}a_i(k|k)\\ b_i(k|k)\end{bmatrix} \tag{A.10a}$$

and

$$F_{i,k+2}(k) = \begin{bmatrix}2(\lambda_i(k|k) - \lambda_{k+2}(k|k)) & 1\\ 1 & 0\end{bmatrix}. \tag{A.10b}$$

Next, $f_{\tilde{z}_k|\tilde{\sigma}_{k-1}}(\tilde{z}_k|\tilde{Z}_{k-1})$ is determined by integrating the pdf in (A.4) with respect to $\tilde{x}_k$. Due to the modular structure of this pdf, the integral is evaluated analytically as

$$f_{\tilde{z}_k|\tilde{\sigma}_{k-1}}(\tilde{z}_k|\tilde{Z}_{k-1}) = \prod_{i=1}^{k+2} \frac{\bar{a}_i(k|k)\tilde{x}_k + \bar{b}_i(k|k)}{\omega_i(k|k)}. \tag{A.11}$$

Finally, the conditional pdf in (A.2) is obtained by dividing the result in (A.4) by the one in (A.11), yielding

$$f_{\tilde{x}_k|\tilde{\sigma}_k}(\tilde{x}_k|\tilde{Z}_k) = \prod_{i=1}^{k+2} \frac{a_i(k|k)\tilde{x}_k + b_i(k|k)}{(\tilde{x}_k - \lambda_i(k|k))^2 + \omega_i^2(k|k)}. \tag{A.12}$$

The terms $a_i(k|k)$ and $b_i(k|k)$ are obtained by dividing $\bar{a}_i(k|k)$ and $\bar{b}_i(k|k)$ by the result in (A.11), i.e., for $i = 1 \cdots, k+2$

$$a_i(k|k) = \frac{\bar{a}_i(k|k)}{f_{\tilde{z}_k|\tilde{\sigma}_{k-1}}(\tilde{z}_k|\tilde{Z}_{k-1})} \tag{A.13}$$

$$b_i(k|k) = \frac{\bar{b}_i(k|k)}{f_{\tilde{z}_k|\tilde{\sigma}_{k-1}}(\tilde{z}_k|\tilde{Z}_{k-1})}. \tag{A.14}$$

(A.12) indicates that the pdf structure defined in (A.1) is maintained in factored form after a measurement update with one additional term in the pdf sum.

It should be pointed out that $f_{\tilde{x}_k|\tilde{\sigma}_k}(\tilde{x}_k|\tilde{Z}_k)$ of (A.12) could be used to compute the state estimate and its estimation error covariance. This result is not shown here, because it is not used in deriving the Cauchy controller. The interested reader is referred to Idan and Speyer (2010) for details.

Next we consider the time propagation from step k to $k+1$. Given the conditional pdf in (A.12), our goal is to construct $f_{\tilde{x}_{k+1}|\tilde{\sigma}_k}(\tilde{x}_{k+1}|\tilde{Z}_k)$, given by the Chapman–Kolmogorov equation

$$f_{\tilde{x}_{k+1}|\tilde{\sigma}_k}(\tilde{x}_{k+1}|\tilde{Z}_k) = \int_{-\infty}^{\infty} f_{\tilde{x}_{k+1}|\tilde{x}_k}(\tilde{x}_{k+1}|\tilde{x}_k)\, f_{\tilde{x}_k|\tilde{\sigma}_k}(\tilde{x}_k|\tilde{Z}_k)\, d\tilde{x}_k, \tag{A.15}$$

where

$$f_{\tilde{x}_{k+1}|\tilde{x}_k}(\tilde{x}_{k+1}|\tilde{x}_k) = \frac{\beta/\pi}{(\tilde{x}_{k+1} - \phi\tilde{x}_k)^2 + \beta^2}. \tag{A.16}$$

Using (A.12) and (A.16), the integral in (A.15) can be computed for each term in the sum analytically, leading to

$$f_{\tilde{x}_{k+1}|\tilde{\sigma}_k}(\tilde{x}_{k+1}|\tilde{Z}_k) = \prod_{i=1}^{k+2} \frac{a_i(k+1|k)\tilde{x}_{k+1} + b_i(k+1|k)}{(\tilde{x}_{k+1} - \lambda_i(k+1|k))^2 + \omega_i^2(k+1|k)}, \tag{A.17}$$

where for $i = 1, \ldots, k+2$

$$\lambda_i(k+1|k) = \phi\lambda_i(k|k) \tag{A.18}$$

$$\omega_i^2(k+1|k) = \phi^2\omega_i^2(k|k) + \beta \tag{A.19}$$

$$\begin{bmatrix}a_i(k+1|k)\\ b_i(k+1|k)\end{bmatrix} = G_i(k)\begin{bmatrix}a_i(k|k)\\ b_i(k|k)\end{bmatrix} \tag{A.20}$$

with

$$G_i(k) = \begin{bmatrix}\operatorname{sign}(\phi) & 0\\ \frac{\lambda_i(k|k)\beta}{\omega_i(k|k)} & \frac{\omega_i(k+1|k)}{\omega_i(k|k)}\end{bmatrix}. \tag{A.21}$$

This shows that the factored pdf structure of (A.1) is regained in (A.17) after a measurement update and a time propagation step.

Appendix B. Gaussian model predictive controller

To compare the performance of the Cauchy controller with one assuming Gaussian noises, one has to consider: (a) the parameters of the Gaussian pdf's that best approximate the Cauchy pdf's, and (b) the objective function used to design the Gaussian controller to be comparable to the one used in the Cauchy noise setting. Those two items are addressed first in this appendix, which concludes with the solution to the Gaussian controller problem.

B.1. Normal pdf least squares fit of a Cauchy pdf

To construct a normal or Gaussian pdf that best fits a given Cauchy pdf, the following optimization problem is solved

$$\xi^* = \arg\min_{\sigma} \int_{-\infty}^{\infty} \left[f_X^C(x) - f_X^N(x)\right]^2 dx, \tag{B.1}$$

where the Cauchy pdf is

$$f_X^C(x) = \frac{\mu/\pi}{x^2 + \mu^2}, \quad \mu > 0 \tag{B.2}$$

and the normal pdf is given by

$$f_X^N(x) = \frac{e^{-x^2/(2\xi^2)}}{\sqrt{2\pi\xi}}, \quad \xi > 0. \tag{B.3}$$

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

Solving the integral in (B.1) analytically and equating its derivative with respect to ξ to zero yields the nonlinear equation

$$\frac{d}{d\xi}\int_{-\infty}^{\infty}\left(\frac{\mu/\pi}{x^2+\mu^2}-\frac{e^{-x^2/(2\xi^2)}}{\sqrt{2\pi}\cdot\xi}\right)dx=\frac{1}{\sqrt{2\pi}\cdot\xi^2}\times\left(1+\kappa^2\left(1-\operatorname{erf}\left(\frac{\kappa}{\sqrt{2}}\right)\right)e^{\frac{\kappa^2}{2}}-\frac{\kappa}{\sqrt{2\pi}}\right)=0, \tag{B.4}$$

where $\kappa=\mu/\xi$. This equation clearly indicates that the optimal ξ is proportional to μ, i.e., $\xi^*=\kappa_0\mu$, where κ_0 is the solution of (B.4). The latter can be solved only numerically to yield $\kappa_0\approx 1.3898$.

Consequently, the equivalent corresponding pdf's for the Gaussian optimal control problem are chosen as

$$f_{X_0}(x_0)=\frac{e^{-(x_0-\bar{x}_0)^2/(2M_0)}}{\sqrt{2\pi M_0}},\quad M_0=\kappa_0^2\sigma^2 \tag{B.5a}$$

$$f_{W_k}(w_k)=\frac{e^{-w_k^2/(2W)}}{\sqrt{2\pi W}},\quad W=\kappa_0^2\beta^2 \tag{B.5b}$$

$$f_{V_k}(v_k)=\frac{e^{-v_k^2/(2V)}}{\sqrt{2\pi V}},\quad V=\kappa_0^2\gamma^2. \tag{B.5c}$$

B.2. Linear-exponential-Gaussian (LEG) m-step optimal model predictive controller

To best approximate the Cauchy objective function of (6) in the Gaussian case, an exponential objective function is chosen. Since only the dynamic programming solution to the LEG problem available (Speyer & Chung, 2008, Chapter 10), the m-step optimal model predictive LEG controller is derived here.

The LEG performance index is

$$J_G^*=\max_{U_k\in F}E\left\{E\left\{e^{-\frac{1}{2}\left(q_p\mathbf{x}_p^2+\sum_{i=k}^{p-1}r_i\mathbf{u}_i^2\right)}\middle|Z_k\right\}\right\}=\max_{U_k\in F}E\left\{e^{-\frac{1}{2}\left(q_p\mathbf{x}_p^2+\sum_{i=k}^{p-1}r_i\mathbf{u}_i^2\right)}\middle|Z_k\right\}$$

$$\text{t:. } E\{J^*_{G_{Z_k}}\}. \tag{B.6}$$

Similarly to fitting the pdf's discussed above, the parameters in the Gaussian objective function can be chosen to best fit the Cauchy objective function as

$$q_p=\frac{1}{\kappa_p^2},\qquad r_i=\frac{1}{\kappa_0^2\zeta_i^2},\quad i=k,\ldots,p-1. \tag{B.7}$$

To solve the maximization problem in (B.6), Z_k could be decomposed as in Section 5. However in this decomposition does not simplify the derivation as was the case when determining the Cauchy controller, and hence is not performed here.

To maximize (B.6) it is sufficient to maximize $J_{G_{\tilde{Z}_k}}$, given by

$$J_{G_{Z_k}}=E\left\{e^{-\frac{1}{2}\left(q_p\mathbf{x}_p^2+\sum_{i=k}^{p-1}r_i\mathbf{u}_i^2\right)}\middle|Z_k\right\}=e^{-\frac{1}{2}\sum_{i=k}^{p-1}r_i\mathbf{u}_i^2}\int_{-\infty}^{\infty}e^{-\frac{1}{2}q_p x_p}f_{X_p|Y_k}(x_p|Z_k)\,dx_p. \tag{B.8}$$

The conditional density in (B.8) is the Gaussian conditional density of the state at time p given all measurements and controls up to time k and is given by

$$f_{\mathbf{x}_p|\sigma_k}(x_p|Z_k)=\frac{1}{\sqrt{2\pi\cdot M_p}}e^{-\frac{1}{2}\frac{(x_p-\bar{\mathbf{x}}_p)^2}{M_p}}, \tag{B.9}$$

where $\bar{\mathbf{x}}_p$ and M_p are the state estimate and error variance at time p of a Kalman filter, generated from the update equations (Speyer & Chung, 2008, Chapter 9)

$$\hat{\mathbf{x}}_k=\bar{\mathbf{x}}_k+P_kH/V(\mathbf{z}_k-H\bar{\mathbf{x}}_k),\quad P_k=\frac{VM_k}{V+H^2M_k}, \tag{B.10a}$$

and propagated to time p from time k using

$$\mathbf{x}_{k+1}=\Phi\hat{\mathbf{x}}_k+\mathbf{u}_k,\qquad M_{k+1}=\Phi^2P_k+W. \tag{B.10b}$$

Using (B.9) in (B.8) and solving the integral gives

$$J_{G_{Z_k}}=\frac{e^{-\frac{1}{2}\left(\frac{q_p\hat{\mathbf{x}}_p^2}{M_pq_p+1}+\sum_{i=k}^{p-1}r_i\mathbf{u}_i^2\right)}}{\sqrt{M_pq_p+1}}, \tag{B.11a}$$

$$\bar{\mathbf{x}}_p=\Phi^m\hat{\mathbf{x}}_k+\sum_{i=k}^{p-1}\Phi^{p-1-i}\mathbf{u}_i. \tag{B.11b}$$

The control input sequence that maximizes this objective function is found analytically. The control input applied at time k is the first element of the optimal control sequence u, found in (B.12), as

$$\mathbf{u}_k^*=-(QS^TS+R)^{-1}S^TQ\Phi^m\hat{\mathbf{x}}_k, \tag{B.12}$$

where

$$S=\begin{bmatrix}\Phi^{m-1}&\cdots&\Phi&1\end{bmatrix},\qquad Q=\frac{q_p}{M_pq_p+1}, \tag{B.13}$$

and R is a diagonal matrix with r_i, $i=1,\ldots,m$ along its diagonal. For the inverse in (B.12) to exist, it is sufficient that $q_p\geq 0$ and that the matrix R be positive definite, which means that $r_i>0\ \forall i$. Note that in the Gaussian case the multiplicative form of the objective function allows a very simple analytical form of the optimal control. The same is true for the dynamic programming solution, where the objective function composed of sums of exponentials appears intractable.

References

Cannon, M., Cheng, Q., & Rakovic, S. V. (2012). Stochastic tube MPC with state estimation. *Automatica 48*, , 536–541.

Dreyfus, S. E. (1965). *Dynamic programming and the calculus of variations*. (p. 211). New York, London: Academic Press.

Fernández, J., Speyer, J.L., & Idan, M. (2013). A stochastic controller for vector linear systems with additive cauchy noises. In *IEEE conference on decision and control*, Firenze, Italy. Dec.

Fernholz, L. T., Morgenthaler, S., & Tukey, J. W. (2004). An outlier nomination method based on the multihalver. *Journal of Statistical Planning and Inference, 122*(1–2), 125–139.

Fleming, W. H., & Rishel, R. W. (1975). *Deterministic and stochastic optimal control*. New York: Springer Verlag (Chapters 5–6).

Fletcher, R., & Powell, M. J. D. (1963). A rapidly convergent descent method for minimization. *The Computer Journal, 6*(2), 163–168.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1968). *Robust statistics*. New York: Wiley (Chapter 1).

Holland, P. W., & Welsch, R. E. (1977). *Communications in Statistics. Theory and Methods, A6*(9), 813–827.

Idan, M., Emadzadeh, A.A., & Speyer, J.L. (2010). Optimal control for a scalar one-step linear system with additive cauchy noise. In *American control conference*, Baltimore, Maryland. June.

Idan, M., & Speyer, J.L. (2008). Cauchy estimation for linear scalar systems. In *IEEE conference on decision and control* (pp. 658–665). Cancun, Mexico, Dec.

Idan, M., & Speyer, J. L. (2010). Cauchy estimation for linear scalar systems. *IEEE Transactions on Automatic Control, 55*(6).

Idan, M., & Speyer, J.L. (2013). Multivariate Cauchy estimator with scalar measurement and process no… Firenze, Italy. Dec.

Jacobson, D. H. (1973). Optimal st… and their relalation to determ… *Automatic Control, 18*(2).

Jun,, & Bitmead, R. R. (2005). Inc… control and its application to n… Q5

Kuruoglu, E. E., Fitzgerald, W. J., & … Rayner, P. J. W. (1998). Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution. *IEEE Communications Letters, 2*(10), 282–284.

Speyer
2013-12-02 22:03:14

Yan, J.

Please cite this article in press as: Speyer, J. L., et al. A stochastic controller for a scalar linear system with additive Cauchy noise. Automatica (2013), http://dx.doi.org/10.1016/j.automatica.2013.11.005

Morari, M., & Zafiriou, E. (1989). *Robust process control*. New Jersey: Prentice-Hall (Chapter 2).
Myers, G. E. (1968). Properties of the conjugate-gradient and Davidon methods. *Journal of Optimization Theory and Applications*, *2*, 209–219. http://dx.doi.org/10.1007/BF00937366.
Pirinen, T. W. (2008). A confidence statistic and an outlier detector for difference estimates in sensor arrays. *IEEE Sensors Journal*, *8*(12), 2008–2015.
Reeves, P. (1969). *A non-gaussian turbulence simulation, Air force flight dynamics laboratory, Tech. Rep. AFFDL-TR-69-67*.
Samorodnitsky, G., & Taqqu, M. S. (1994). *Stable non-Gaussian random processes: stochastic models with infinite variance*. New York: Chapman & Hall.
Speyer, J. L., & Chung, W. H. (2008). *Stochastic processes, estimation, and control* (1st ed.). SIAM.
Speyer, J.L., Idan, M., & Fernández, J. (2010). Multi-step prediction optimal control for a scalar linear system with additive cauchy noise. In *IEEE conference on decision and control*, Atlanta, Georgia. Dec.
Tsakalides, P., & Nikias, C. L. (1998). Deviation from normality in statistical signal processing: parameter estimation with alpha-stable distributions. In *A practical guide to heavy tails: statistical techniques and applications* (pp. 379–404). Boston: Birkhauser.

Jason L. Speyer received the B.S. in Aeronautics and Astronautics from MIT, and the Ph.D. in applied mathematics from Harvard University. He is the Ronald and Valerie Sugar Distinguished Professor in Engineering within the Mechanical and Aerospace Engineering Department and the Electrical Engineering Department, UCLA. He coauthored, with W.H. Chung, *Stochastic Processes, Estimation, and Control* (SIAM, 2008), and coauthored, with D.H. Jacobson, *Primer on Optimal Control Theory* (SIAM, 2010). He is a life fellow of the IEEE and a fellow of the AIAA and was awarded the AIAA Mechanics and Control of Flight Award, AIAA Dryden Lectureship in Research, Air Force Exceptional Civilian Decoration (1991 and 2001), IEEE Third Millennium Medal, the AIAA Aerospace Guidance, Navigation, and Control Award, and membership in the National Academy of Engineering.

Moshe Idan received his B.Sc. and M.Sc. degrees from the Faculty of Aerospace Engineering at the Technion in 1983 and 1986, respectively. He earned his Ph.D. from the Department of Aerospace Engineering at Stanford University in 1990. Since 1991, Dr. Idan has been with the Faculty of Aerospace Engineering at the Technion. In 2000–2001 he was a visiting research scholar at the School of Aerospace Engineering at Georgia Tech, and in 2007–2009 he spent a two year sabbatical at the Department of Mechanical and Aerospace Engineering in UCLA. His current research interests include robust and adaptive flight control system design techniques and applications, innovative future air traffic control concepts, and control and estimation of systems with non-Gaussian noises.

Javier H. Fernandez received his B.S. in Engineering Sciences from Dartmouth College in Hanover, New Hampshire, in 2008, and his M.S. degree in Mechanical Engineering from the University of California, Los Angeles, in 2010. He continues there, pursuing his Ph.D. in the area of control theory. His current interests include stochastic optimal control with non-classical information patterns.

E Appendix

Javier H. Fernández, Jason L. Speyer, and Moshe Idan
A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises
Precedings of the 2013 IEEE Conference on Decision and Control

A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises

Javier Fernández
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: jhf@seas.ucla.edu

Jason L. Speyer
Mechanical and Aerospace Engineering
University of California, Los Angeles
Los Angeles, California 90095–1597
Email: speyer@seas.ucla.edu

Moshe Idan
Aerospace Engineering
Technion - Israel Institute of Technology
Haifa, Israel, 32000
Email: moshe.idan@technion.ac.il

***Abstract*—An optimal predictive controller for linear, vector-state dynamic systems driven by Cauchy measurement and process noises is developed. For the vector-state system, the probability distribution function (pdf) of the state conditioned on the measurement history cannot be generated. However, the characteristic function of this pdf can be expressed in an analytic form. Consequently, the performance index is evaluated in the spectral domain using this characteristic function. By using an objective function that is a product of functions resembling Cauchy pdfs, the conditional performance index is obtained analytically in closed form by using Parseval's equation and integrating over the spectral vector. This forms a non-convex function of the control signal, and must be optimized numerically at each time step. A two-state example is used to expose the interesting robustness characteristics of the proposed controller.**

I. Introduction

Models in modern stochastic optimal control algorithms like the linear quadratic Gaussian (LQG) and the linear exponential Gaussian (LEG) assume linear dynamics and additive process and measurement noises described by the Gaussian probability density function (pdf). The Gaussian distribution function has very light tails, so that large deviations are essentially impossible. Therefore, the LQG and LEG algorithms do not perform well in the presence of heavy-tailed or impulsive uncertainties. In many practical applications, such as radar and sonar systems affected by atmospheric and underwater acoustic noises, more impulsive uncertainties are observed [1]. Impulsive behavior is also more effective at modeling adversarial motion, as is air turbulence, which is better described by distributions with heavier tails than the Gaussian [2].

Therefore, in this paper we propose a system model that assumes linear dynamics driven by additive process and measurement noises described by Cauchy pdfs. Both the Cauchy and Gaussian pdfs belong to a class of distributions called the symmetric α-stable (Sα-S) class, whose members are described by their characteristic functions. A full treatment of the Sα-S class can be found in [3]. The Cauchy pdf is in a subset of this class whose members have infinite second moments. In addition, the mean of the Cauchy pdf is not well defined.

Algorithms for optimal estimation and control of scalar linear systems driven by Cauchy distributed process and measurement noises have been developed previously in [4, 5]. There, the conditional performance index for model predictive control is determined directly by taking the conditional expectation of the objective function using the probability density given the measurement history as presented in [4]. A dynamic programming algorithm is also developed in [5]. It is shown that the solution to the dynamic programming recursion is intractable because of the need to average over future measurements in determining the optimal return function.

In this paper, the Cauchy optimal control algorithm for scalar systems [5] is extended to systems with a vector state. For the vector case, the conditional pdf (cpdf) given the measurement history is not available. However the characteristic function of the cpdf can be recursively propagated [6, 7]. The significant contribution in this paper is evaluating, in closed form, the conditional performance index using the cpdf's characteristic function instead of the cpdf itself, and integrating over the spectral variables instead of the state variables.

Although the cpdf is not available as a function of the state vector, the conditional expectation of the objective function, i.e. the conditional performance index, can be computed using the characteristic function of the cpdf, which is available as a function of the spectral vector [7]. The objective function is cast as a product of functions resembling Cauchy pdfs, which are easily transformed into a function of the spectral variables. Consequently, the conditional performance index is found in a closed form. Due to its complexity, the optimal control signal is determined by numerically optimizing this conditional performance index in a model predictive control setting.

The remainder of the paper is structured as follows. The controlled system model is presented in Section II. An appropriate, computable performance index for this problem is presented in Section III and subsequently transformed from the state variable to the spectral variable form. In Section IV the spectral integrations required to determine the conditional performance index are reduced to an integral formula that can be evaluated in closed form. Section V addresses a special case of systems with two states. Here, using an alternative, simplified form of the two state cpdf's characteristic function, the conditional performance is determined in closed form. In Section VI numerical examples are given. Conclusions are given in Section VII.

This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-10-1-0570, and by the United States - Israel Binational Science Foundation, Grant 2008040.

II. Description of the Model

This paper deals with a discrete time, linear system described by

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Lambda u(k) + \Gamma w(k) \\ z(k) &= \mathrm{H} x(k) + v(k) \end{aligned} \tag{1}$$

where $x(k) \in \mathrm{R}^n$ is the state vector, $u(k)$ is a scalar deterministic input, $z(k)$ is a scalar measurement, and $w(k)$ and $v(k)$ are scalar independent Cauchy distributed process and measurement noise inputs with medians at zero and scaling parameters of β and γ, respectively, so that their pdfs are given by

$$f_{W_k}(w(k)) = \frac{\beta/\pi}{w^2(k) + \beta^2}, \quad f_{V_k}(v(k)) = \frac{\gamma/\pi}{v^2(k) + \gamma^2}. \tag{2}$$

The characteristic functions of these pdfs are

$$\varphi_{W_k}(\sigma) = e^{-\beta|\sigma|}, \quad \varphi_{V_k}(\sigma) = e^{-\gamma|\sigma|}, \tag{3}$$

where σ is the scalar spectral variable.

The initial conditions are assumed to be independent Cauchy distributed random variables with the pdfs

$$f_{X_0}(x(0)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{(x_i(0) - \bar{x}_i(0))^2 + \alpha_i^2}. \tag{4a}$$

Its characteristic function is given by

$$\varphi_{X_0}(v) = \prod_{i=1}^{n} e^{-\alpha_i |v_i| + j\bar{x}_i(0) v_i}. \tag{4b}$$

where v_i is an element of $v \in \mathrm{R}^n$.

The stochastic system (1) can be decomposed into two systems, one driven by $u(k)$ and one by $w(k)$, by exploiting the linearity of the system. Let $\bar{x}(k)$ and $\bar{z}(k)$ be the part of the system driven by the control $u(k)$ only, and $\tilde{x}(k)$ and $\tilde{z}(k)$ be the part of the system driven by the process noise $w(k)$ only and contains all the underlying random variables. Then,

$$x(k) = \bar{x}(k) + \tilde{x}(k) \tag{5a}$$
$$z(k) = \bar{z}(k) + \tilde{z}(k). \tag{5b}$$

The controlled part of the system is described by

$$\bar{x}(k) = \Phi \bar{x}(k-1) + \Lambda u(k-1) \tag{6a}$$
$$\bar{z}(k) = \mathrm{H} \bar{x}(k) \tag{6b}$$

with initial condition $\bar{x}(0)$. The process noise driven part is given by

$$\tilde{x}(k) = \Phi \tilde{x}(k-1) + \Gamma w(k-1) \tag{7a}$$
$$\tilde{z}(k) = \mathrm{H} \tilde{x}(k) + v(k). \tag{7b}$$

The process and measurements noise pdfs were defined in (2), while the initial condition of this stochastic model is Cauchy distributed with a pdf given by

$$f_{\tilde{X}_0}(\tilde{x}(0)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{\tilde{x}_i^2(0) + \alpha_i^2}. \tag{8a}$$

Its characteristic function is

$$\varphi_{\tilde{X}_0}(v) = \prod_{i=1}^{n} e^{-\alpha_i |v_i|}. \tag{8b}$$

The above decomposition will be used to derive the Cauchy controller.

Let the state, measurement, and control histories used in the control problem formulation be defined as

$$X_R^m := \{x(R), \ldots, x(m)\}, \tag{9a}$$
$$\tilde{Z}_k := \{\tilde{z}(0), \ldots, \tilde{z}(k)\}, \tag{9b}$$
$$U_R^m, \; U_R^m \in F \tag{9c}$$

where F is the class of piecewise continuous functions adapted to the σ-algebra σ_k generated by the measurement history, i.e. the control is a random variable that is measurable with respect to events in σ_k [8]. Moreover, in [9] it is shown that u_k is adapted to the σ-algebra $\tilde{\sigma}_k$ generated by $\tilde{Z}_k$, which means that the control is measurable on events generated by $\tilde{Z}_k$ only.

III. Derivation of the Cost Using Characteristic Functions

Our proposed controller is an m-step model predictive controller [10] that uses current and past measurements, and averages over future process noise. At each time step, the conditional performance index is computed. Since the performance index will be shown to be a nonconvex function of the control sequence, it is maximized numerically. Once the optimal control sequence of length m is computed, only the first control in that sequence is applied. At the next step, a new measurement is taken and the process is repeated, producing a new optimal control sequence and applying only the first one. In this paper, we study the optimal stochastic state regulation problem, noting that the tracking problem can be handled in a similar fashion. Our regulation problem will have a finite horizon of length m such that the terminal state occurs at time-step $p = k + m$.

Similar to the scalar control problem presented in [5], the control objective function is chosen as a product of Cauchy-like functions given by

$$\psi\left(X_{k+1}^p, U_k^{p-1}\right) = \prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{u^2(i) + \zeta_i^2} \cdot \prod_{r=1}^{m} \frac{\eta_{i+1,r}/\pi}{x_r^2(i+1) + \eta_{i+1,r}^2}. \tag{10}$$

Then, the the performance index conditioned on the current measurement history and averaged over future process noises is given by

$$\begin{aligned} J_{k,p}^* &= \max_{U_k^{p-1} \in F} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right)\right] \\ &= \max_{U_k^{p-1} \in F} E\left[E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \middle| \tilde{Z}_k\right]\right] \\ &= E\left[\max_{U_k^{p-1} \in F} E\left[\psi\left(X_{k+1}^p, U_k^{p-1}\right) \middle| \tilde{Z}_k\right]\right] = E\left[J_{\tilde{Z}_k}^*\right], \end{aligned} \tag{11}$$

where the interchange of the maximum and expectation operations is due to the fundamental theorem in [11].

We are now concerned with determining the analytic form for the conditional performance index $J_{\tilde{Z}_k}^*$. Using (10), it

becomes

$$J_{Z_k} = E\{\psi(X_{k+1}^p, U_k^{p-1}) \mid Z_k\}$$
$$= E\Bigg\{\prod_{i=k}^{p-1}\prod_{r=1}^{n}\frac{\eta_{i+1,r}/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2}\cdot\frac{\zeta_i/\pi}{u_i^2(i)+\zeta_i^2}\Bigg\}$$
$$= \int_{-\infty}^{\infty}\prod_{i=k}^{p-1}\prod_{r=1}^{n}\frac{\eta_{i+1,r}/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2}\cdot\frac{\zeta_i/\pi}{u_i^2(i)+\zeta_i^2}$$
$$\times f_W(\tilde{x}(p)|\tilde{x}(p-1))\cdots f_W(\tilde{x}(k+2)|\tilde{x}(k+1))$$
$$\times f_{\tilde{X}_{k+1}|Z_k}(\tilde{x}(k+1)|Z_k)\,d\tilde{x}_1(k+1)\ldots d\tilde{x}_n(k+1)$$
$$\times d\tilde{x}_1(k+2)\ldots d\tilde{x}_n(k+2)\ldots d\tilde{x}_1(p)\ldots d\tilde{x}_n(p) \quad (12)$$

For now, let us only consider weighting on the terminal state, $x(p)$, and on the m scalar control inputs. The control weighting functions $M^U\prod_{i=k}^{p-1}\frac{\zeta_i/\pi}{u_i^2(i)+\zeta_i^2}$ can come out of the integral. Then, the product over i inside the integrand in (12) has only the term for $i = p-1$, and the integral is only over $\{\tilde{x}_1(p),\ldots,\tilde{x}_n(p)\}$. Thus, for notational convenience we can drop the time-step index in this integral and write it over $\{\tilde{x}_1,\ldots,\tilde{x}_n\}$ as

$$J_{Z_k} = M^U\int_{-\infty}^{\infty}\prod_{r=1}^{n}\frac{\eta_{p,r}/\pi}{\tilde{x}_r+\bar{x}_r{}^2+\eta_{p,r}^2}$$
$$\times f_{\tilde{X}_p|Z_k}(\tilde{x}|Z_k)\,d\tilde{x}_1\ldots d\tilde{x}_n. \quad (13)$$

The cpdf $f_{\tilde{X}_p|Z_k}(\tilde{x}|Z_k)$ can be evaluated in closed form for scalar systems [4]. However, for vector state systems it is the characteristic function of the cpdf, $\varphi_{\tilde{X}_p|Z_k}(v)$, that is evaluated in closed form. Therefore, when computing the conditional performance index, we need to be able to integrate over the spectral variable v instead of the pdf variable $\tilde{x}$.

Define the product over r in the integral in (13) as R_x and its Fourier transform as L_x,

$$R_x(\tilde{x}+\bar{x}) = \prod_{r=1}^{n}\frac{\eta_{i+1,r}/\pi}{\tilde{x}_r+\bar{x}_r{}^2+\eta_{i+1,r}^2} \quad (14a)$$
$$L_x(v) = \prod_{r=1}^{n}e^{-\eta_{p,r}|v_r|-j\bar{x}_r v_r}. \quad (14b)$$

Using these definitions, we can apply Parseval's equation over each variable in (13) to express the conditional performance index as an integral over the spectral variable v,

$$J_{Z_k} = \frac{M^U}{(2\pi)^n}\int_{-\infty}^{\infty}L_x^*(v)\cdot\varphi_{\tilde{X}_p|Z_k}(v)\,dv\ldots dv$$
$$= \frac{M^U}{(2\pi)^n}\int_{-\infty}^{\infty}\prod_{r=1}^{n}e^{-\eta_{p,r}|v_r|+j\bar{x}_r(p)v_r}$$
$$\times\varphi_{\tilde{X}_p|Z_k}(v)\,dv_1\ldots dv_n, \quad (15)$$

where L_x^* is the complex conjugate of L_x. The next section

IV. The Conditional Performance Index

Consider the integral over v_n in (15),

$$I_n = \int_{-\infty}^{\infty}\prod_{r=1}^{n}e^{-\eta_{p,r}E_{r,v}|}e^{j\bar{x}(p),v}\varphi_{\tilde{X}_p|Z_k}(v)dv_n$$
$$= \int_{-\infty}^{\infty}e^{-\sum_{r=1}^{n}\eta_{p,r}|E_{r,v}|+j\bar{x}(p),v}\varphi_{\tilde{X}_p|Z_k}(v)dv_n \quad (16)$$

The cpdf for the state $\tilde{x}(k)$ is denoted as $f_{\tilde{X}_k|Z_k}$. The unnormalized cpdf (ucpdf) is denoted as $\bar{f}_{\tilde{X}_k|Z_k} = f_{\tilde{X}_k|Z_k}\cdot f_{Z_k}$ where f_{Z_k} is the pdf of the measurement history and has a known value. In [12, 13], the characteristic function of the ucpdf $\bar{\varphi}_{\tilde{X}_k|Z_k}(v)$ is recursively propagated; the characteristic function of the normalized cpdf is $\varphi_{\tilde{X}_k|Z_k}(v) = \bar{\varphi}_{X_k|Z_k}(v)/f_{Z_k}$, where $f_{Z_k} = \varphi_{\tilde{X}_k|Z_k}(v)|_{v=0}$. From [12, 13] the form of the characteristic function of the ucpdf at time k is shown to be

$$\bar{\varphi}_{\tilde{X}_k|Z_k} = \sum_{i=1}^{n_t^{k|k}}g_i^{k|k}(y_{gi}(v))e^{y_{ei}^{k|k}(v)} \quad (17a)$$

where

$$y_{gi}^{k|k}(v) = \sum_{R=1}^{n_{ei}^{k|k}}q_{i,R}^{k|k}\operatorname{sgn}\left(a_{iR}^{k|k},v\right)\in\mathbb{R}^k \quad (17b)$$

$$v_{ei}^{k|k}(v) = -\sum_{R=1}^{n_{ei}^{k|k}}p_{i,R}^{k|k}\left(a_{iR}^{k|k},v\right)+j\left(b_i^{k|k},v\right) \quad (17c)$$

and the parameters $n_t^{k|k}$, $n_{e,i}^{k|k}$, $q_{iR}^{k|k}$, $p_{iR}^{k|k}$, $a_{iR}^{k|k}$, $b_i^{k|k}$ are generated sequentially from $k = 0$.

For the MPC algorithm, the characteristic function of the ucpdf is to be propagated through the stochastic dynamics to time $k+m = p$. This characteristic function $\bar{\varphi}_{\tilde{X}_p|Z_k}(v)$ is

$$\bar{\varphi}_{\tilde{X}_p|Z_k}(v) = \bar{\varphi}_{\tilde{X}_k|Z_k}(\Phi^{m\mathsf{T}}v)\varphi_W((\Phi^{m-1}\Gamma)^{\mathsf{T}}v)$$
$$\times\cdots\times\varphi_W((\Phi\Gamma)^{\mathsf{T}}v)\varphi_W(\Gamma v)$$
$$= \sum_{i=1}^{n_t^{k|k}}g_i^{k|k}(y_{gi}(\Phi^{m\mathsf{T}}v))e^{y_{ei}^{k|k}(\Phi^{m\mathsf{T}}v)}$$
$$\times\exp\left(-\beta\left(\Phi^{m-1}\Gamma,v\right)-\cdots-\beta\left(|\Phi\Gamma,v)\right|-\beta\left|(\Gamma,v)\right|\right) \quad (18)$$

In (18) we add m terms to the sum in $y_{ei}^{k|k}(v)$ of (17c). By combining the exponent in (16) with that in (18), the combined exponent in the integrand of (16) has in total of $n_{ei}^{k|k}+m+n$ real terms, and the imaginary part is composed of two components. Define the following terms

$$\bar{p}_{iR} = p_{iR},\quad \bar{a}_{iR}^{k|k} = \Phi^m a_{iR}^{k|k}\quad \text{for } R = 1,\ldots,n_{ei}^{k|k}$$
$$\bar{p}_{iR} = \beta,\quad \bar{a}_{iR}^{k|k} = \Phi^t\Gamma\quad \text{for } {\scriptstyle R=1,\ldots,n_{ei}^{k|k}\atop\scriptstyle t=R-(n_{ei}^{k|k}+1)} \quad (19a)$$
$$\bar{p}_{iR} = \eta_{pR},\quad \bar{a}_{iR}^{k|k} = {}_r\quad \text{for } {\scriptstyle R=1,\ldots,n_{ei}^{k|k}\atop\scriptstyle r=R-(n_{ei}^{k|k}+m)}$$

and

$$\bar{b}^{k|k}\quad {}^{k|k}\quad {}^{k|k}\quad {}^{k|k}$$

shows how to evaluate these n nested integrals sequentially in closed form.

$$\bar{x}_i = b_{k|k} + \bar{x}_p, \qquad \bar{n}_{ei} = n_{k|k} + m + n_{k|k} \qquad (19b)$$
$$q_{tR} = 0 \qquad \text{for } R = n_{ei} + 1, \ldots, \bar{n}_{ei}$$

Using these definitions, the integrand in (16) becomes

$$\psi(v) = \prod_{i=1}^{n_t^{k|k}} \psi_i(v) = \prod_{i=1}^{n_t^{k|k}} g_i^{k|k}(\bar{y}_{gi}^{k|k}(v)) \cdot e^{\bar{y}_{ei}^{k|k}(v)} \quad (20a)$$

where

$$\bar{y}_{gi}^{k|k}(v) = \sum_{R=1}^{\bar{n}_{gi}^{k|k}} q_{iR}^{k|k} \operatorname{sgn}\left(\left\langle \bar{a}_{iR}^{k|k}, v \right\rangle\right) \quad (20b)$$

$$\bar{y}_{ei}^{k|k}(v) = \sum_{R=1}^{\bar{n}_{ei}^{k|k}} \bar{p}_{iR}^{k|k} \left\langle \bar{a}_{iR}^{k|k}, v \right\rangle + j \left\langle \bar{b}_{iR}^{k|k}, v \right\rangle \quad (20c)$$

The integration in (16) is performed for each element of v in turn. Beginning with v_n, decompose $v = \begin{bmatrix} \hat{v} \\ v_n \end{bmatrix}$ where $\hat{v} \in \mathbb{R}^{n-1}$. Then,

$$I_n = \int_{-\infty}^{\infty} \psi(v)dv = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n_t^{k|k}} \psi_i(v)dv$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n_t^{k|k}} \left[\int_{-\infty}^{\infty} \psi_i(v_n, \hat{v})dv_n \right] d\hat{v} \quad (21)$$

The objective is to reduce the inner integral in (21) to a form that is obtained in closed form using the integral formula developed in [12, 13]. First, since $\bar{a}_{iR}^{k|k}$ multiplies v in the sign function in (20b) and the absolute value function in (20c), they are decomposed as $\bar{a}_{iR}^{k|k} = \hat{a}_{iR}^{k|k} \tilde{a}_{iR}^{k|k}$, where $\tilde{a}_{iR}^{k|k}$ is a scalar and $\hat{a}_{iR}^{k|k} \in \mathbb{R}^{n-1}$. Therefore, the inner products in (20b) and (20c) become

$$\left\langle \bar{a}_{iR}^{k|k}, v \right\rangle = \left\langle \hat{a}_{iR}^{k|k}, \hat{v} \right\rangle - \tilde{a}_{iR}^{k|k} v_n \quad . \quad (22)$$

In order to rewrite (22) in a form consistent with the integral formula in [12, 13], $-\tilde{a}_{iR}^{k|k}$ is divided out of the second term. If $\tilde{a}_{iR}^{k|k} = 0$, then the term $e^{\left\langle \hat{a}_{iR}^{k|k}, \hat{v} \right\rangle}$ loses dependence on v_n and it is removed from the inner integral in (21). Therefore, only $\tilde{a}_{iR}^{k|k} \neq 0$ needs to be considered. Let (22) be rewritten as

$$\left\langle \bar{a}_{iR}^{k|k}, v \right\rangle = \tilde{a}_{iR}^{k|k} \operatorname{sgn}\left(-\tilde{a}_{iR}^{k|k}\right) \left(\mu_{iR}^{k|k} - v_n \right) \quad (23a)$$

where $\mu_{iR}^{k|k} = \dfrac{\left\langle \hat{a}_{iR}^{k|k}, \hat{v} \right\rangle}{-\tilde{a}_{iR}^{k|k}}$. Therefore, the elements in (20b) and (20c) are

$$q_{iR}^{k|k} \operatorname{sgn}\left(\left\langle \bar{a}_{iR}^{k|k}, v \right\rangle\right) = \bar{q}_{iR}^{k|k} \operatorname{sgn}\left(\mu_{iR}^{k|k} - v_n\right) \quad (23b)$$

$$\bar{p}_{iR}^{k|k} \left\langle \bar{a}_{iR}^{k|k}, \hat{v} \right\rangle = \rho_{iR}^{k|k} \left(\mu_{iR}^{k|k} - v_n \right) \quad (23c)$$

where $\bar{q}_{iR}^{k|k} = q_{iR}^{k|k} \operatorname{sgn}\left(-\tilde{a}_{iR}^{k|k}\right)$ and $\rho_{iR}^{k|k} = p_{iR}^{k|k} \tilde{a}_{iR}^{k|k}$. Using these definitions, the inner integral of (21) is of the form

$$\int_{-\infty}^{\infty} \psi_i(v)dv_n = e^{j\left\langle \hat{b}_i^{k|k}, \hat{v} \right\rangle} \int_{-\infty}^{\infty} g_i^{k|k}\left[\sum_{R=1}^{\bar{n}_{ei}^{k|k}} \bar{q}_{iR}^{k|k} \operatorname{sgn}\left(\mu_{iR}^{k|k} - v_n\right) \right]$$

$$\times \exp\left[\sum_{R=1}^{\bar{n}_{ei}^{k|k}} \rho_{iR}^{k|k} \left(\mu_{iR}^{k|k} - v_n \right) + j\tilde{b}_i^{k|k} v_n \right] dv_n. \quad (24)$$

The convolution integral in (24) is shown in [12, 13] to have a closed form solution composed as a sum with $\bar{n}_{iR}^{k|k}$ terms, each of which is structurally similar to the terms in $\psi_i(v)$. That is, there will be a new g function which is a function of signs of inner products of $\hat{v}$.

Therefore, this integration process can be repeated until all of the integrals are taken, and a closed form solution of the conditional performance index is determined. This pattern will be seen in the next section, where the conditional performance index for the two state system is explicitly obtained.

V. The Conditional Performance Index for a Second Order System

Now, let us limit our discussion to a second order system in order to use the structure for the cpdf's characteristic function presented in [7]. We use this alternate, two state structure for the characteristic function of the cpdf in order to make the subsequent derivations and computation more tractable. This is due to a simpler structure for the exponential argument in (17c), which produces fewer terms in the sum, as well as a simpler, closed form representation for the $g_i^{k|k}$ coefficients in (17a).

The structure for the characteristic function for the ucpdf is given by

$$\bar{\varphi}_{x_p|Z_k}(v) = \sum_{i=1}^{N_k} G_i(v) \times \exp\left[-\sum_{R=1}^{L_k^i} P_{k,i}^R B_k^{M_{k,i}^R} v + j \sum_{R=1}^{2} Z_{k,i}^R B_k^{M_{k,i}^R} v \right], \quad (25)$$

which consists of a sum of N_k similar terms. Each of these terms has a coefficient $G_i(v)$ and an exponential whose argument involves a sum of absolute values equivalent to (17c). In the exponential argument: there is a sum of L_k^i absolute value terms; the $P_{k,i}^R$ and $Z_{k,i}^R$ values represent real, scalar constants where $P_{k,i}^R > 0$; the $B_k^{M_{k,i}^R}$ s are 1×2 row vectors, called the fundamental directions; and the $M_{k,i}^R$ represent integers that index the fundamental directions that multiply v.

All of these parameters correspond with those of the structure for the cpdf's characteristic function presented in Section IV. The $P_{k,i}$ and $Z_{k,i}$ parameters correspond to $p_{iR}^{k|k}$ and $b_i^{k|k}$, respectively, and the fundamental directions $B_k^{M_i^R}$ are the same as the $a_{iR}^{k|k}$ vectors.

The coefficients $G_i(v)$ equal the $g_i^{k|k}$ from (17a), and are rational polynomial of sums of sign functions, given by

$$G_i(v) = \frac{1}{(2\pi)^n} \left[a_i + b_i \operatorname{sgn}\left(B_k^{M_i^1} v\right) \operatorname{sgn}\left(B_k^{M_i^2} v\right) + jc_i \operatorname{sgn}\left(B_k^{M_i^1} v\right) + jd_i \operatorname{sgn}\left(B_k^{M_i^2} v\right) \right]$$

$$\times \prod_{R=1}^{L_k^i - 2} \frac{1}{j\theta_{k-R}^i + \gamma^i + S_{k-R}^i(B_k v)} - \frac{1}{j\theta_{k-R}^i + \gamma^i + S_{k-R}^i(B_k v)}, \quad (26a)$$

where

$$S^i_{k-R}(B_k v) = \sum_{r=1}^{L^i_k - R - 2} P_i^R \; B_{k-r}^{M_i^R} \mathrm{H}^{\mathrm{T}} \; \mathrm{sgn}\left[B_k^{M_i^R} v \right], \quad (26b)$$

$$\theta^i_{k-R} = z_{k-R} - Z_k B_{k-R}^{M_i^1} \mathrm{H}^{\mathrm{T}} - Z_k^2 B_{k-R}^{M_i^2} \mathrm{H}^{\mathrm{T}}. \quad (26c)$$

The arguments $y^{k|k}_{gi}(v)$ of $g^{k|k}_i$, given in (17b), correspond to the $S^i_{k-R}(B_k v)$ defined above.

Next, the transformed objective function in the performance index is given by

$$L^*_{x(p)}(v) = e^{-\eta_{p,1}|v_1| + j\bar{x}_1 v_1 - \eta_{p2}|v_2| + j\bar{x}_2 v_2} \quad (27)$$

where, since k will be a constant through this process, the time subscript of $\bar{x}(p)$ is replaced with element subscripts as $\bar{x}(p) = [\bar{x}_1 \;\; \bar{x}_2]$. The 2-state specific version of (15) then becomes

$$J_{\tilde{Z}_k} = \frac{M_U}{(2\pi)^2} \int_{-\infty}^{\infty} L^*_{x(p)}(v_1, v_2) \cdot \varphi_{X_p|Z_k}(v_1, v_2)\, dv_1 v_2$$

$$= \frac{M_U}{(2\pi)^2} \int_{-\infty}^{\infty} \exp\left(-\eta_1 |v_1| - \eta_2 |v_2| + j\bar{x}_1 v_1 + j\bar{x}_2 v_2 \right)$$

$$\times \prod_{i=1}^{N_k} G_i(v) \exp\left[-\sum_{R=1}^{L^i_k} P^R_{k,i} \left| B_k^{M^R_{k,i}} v \right| + j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_k v \right] dv. \quad (28)$$

The integral with respect to v_2 is now taken. We use a second subscript to denote the individual elements of the rows of B_k as $B_k^{M_i^R} = \begin{bmatrix} B_{k,1}^{M_i^R} & B_{k,2}^{M_i^R} \end{bmatrix}$. This allows us to decompose the complex part of the exponential argument as

$$j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_k v = j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_{k,1} v_1 + j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_{k,2} v_2, \quad (29)$$

and rewrite $J_{\tilde{Z}_k}$ as

$$J_{\tilde{Z}_k} = \frac{1}{(2\pi)^2} \prod_{i=k}^{Z_k} \int_{p-1}^{\zeta_i/\pi} \frac{1}{u^2(i) + \zeta_i^2}$$

$$\times \prod_{i=1}^{N_k} \int_{-\infty}^{\infty} \exp\left[-\eta_1 |v_1| + j\bar{x}_1 v_1 + j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_{k,1} v_1 \right]$$

$$\times \int_{-\infty}^{\infty} G_i(v) \cdot \exp\left[-\sum_{R=1}^{L^i_k} P^R_{k,i} \left| B^{M^R_{k,i}}_k v \right| - \eta_2 |v_2| \right]$$

$$+ j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_{k,2} v_2 + j\bar{x}_2 v_2 \; dv_2 dv_1. \quad (30)$$

Denote the inner integral with respect to v_2 in (30) as I_2:

$$I_2 = \int_{-\infty}^{\infty} G_i(v_1, v_2)$$

$$\times \exp\left[-\sum_{R=1}^{L^i_k} P^R_{k,i} \left| B^{M^R_{k,i}}_{k,1} v_1 + B^{M^R_{k,i}}_{k,2} v_2 \right| - \eta_2 |0 + v_2| \right.$$

$$\left. + j \sum_{R=1}^{L^i_k} Z^R_{k,i} B^{M^R_{k,i}}_{k,2} v_2 + j\bar{x}_2 v_2 \right] dv_2. \quad (31)$$

The integral I_2 is over the v_2 variable, but it also contains v_1 in the absolute value terms in the argument of the exponential. In order to solve it, we need to use the integral of absolute values method presented in [7], which requires writing the integral in the form given in

$$I_2 = \int_{-\infty}^{\infty} \breve{G}_i \left[\prod_{R=1}^{L^i_k + 1} \breve{\rho}_R \right] \mathrm{sgn}(\mu_R - \sigma)$$

$$\times \exp\left[-\sum_{R=1}^{L^i_k + 1} \breve{\rho}_R |\mu_R - \sigma| + j\xi_k \sigma \right] d\sigma, \quad (32)$$

as was done in obtaining (24). That method involves defining a set of scalar constants ρ_R and ξ^i_k, as well as scalar variables μ_R that depend on v_1, and thus are constants in this integration. In this derivation, we assume that all the $B^{M_i^R}_{k,2} \neq 0$. Then, we can construct a $\{\mu_R\}_1^{L^i_k+1}$ set that transforms I_2 into the integral of absolute values structure. This assumption does not affect generality, because if one of the $B^{M_i^R}_{k,2}$ does equal zero for some i, then it would multiply the v_2 variable by 0 and thus, that absolute value term would not be a function of v_2 and would come out of I_2 and integrated later.

The set of variables, denoted $\{\mu_R\}$, is constructed as

$$B^{M_i^R}_{k,1} v_1 + B^{M_i^R}_{k,2} v_2 = -B^{M_i^R}_{k,2} \left[\frac{-B^{M_i^R}_{k,1} v_1}{B^{M_i^R}_{k,2}} - v_2 \right]$$

$$= -B^{M_i^R}_{k,2} (\mu_R - v_2) \quad (33a)$$

where

$$\mu_R = \begin{cases} \dfrac{-B^{M_i^R}_{k,1} v}{B^{M_i^R}_{k,2}} & R \in \{1, \ldots, L^i_k\}, i \in \{1, \ldots, N_{k-1}\} \\ 0 & R = L^i_k + 1. \end{cases} \quad (33b)$$

Similarly, for the argument of the exponential, we can construct a set of $\{\rho_R\}_1^{L^i_k+1}$ as

$$\rho_R = \begin{cases} P^i_k - B^{M_{iR}}_{k,2} & R \in \{1, \ldots, L^i_k\}, i \in \{1, \ldots, N_{k-1}\} \\ \eta_2 & R = L_k + 1 \end{cases} \quad (33c)$$

and a scalar number ξ^i_k as

$$\xi_k = \bar{x}_2 + \sum_{R=1}^{L^i_k} Z_i B^{M_i^R}_{k,2}. \quad (33d)$$

The solution to an integral of an exponent of absolute values requires dividing the domain of integration into regions in which the integrand is continuous. Since $G_i(v)$ is piecewise-constant, its discontinuities lie on the boundaries of these

regions, and hence $G_i(v)$ is treated as a constant in each integral. In order for the $G_i(v)$ coefficients to be consistent with the form in (32), use the tilde, bar, and hat substitutions similar to the measurement update process described in [7] in order to write G_i in (31) as $\breve{G}_i$:

$$G_i(v_1, v_2) \to \breve{G}_i \Big(\sum_{R=1}^{L_k^i+1} \tilde{\rho}_R \operatorname{sgn}(\mu_R - \sigma), \sum_{R=1}^{L_k^i+1} \bar{\rho}_R \operatorname{sgn}(\mu_R - \sigma), \sum_{R=1}^{L_k^i+1} \hat{\rho}_R^{(L_k^i-2)} \operatorname{sgn}(\mu_R - \sigma), \ldots, \sum_{R=1}^{L_k^i+1} \hat{\rho}_R^{(1)} \operatorname{sgn}(\mu_R - \sigma) \Big) \to \breve{G}_i \Big(\sum_{R=1}^{L_k^i+1} \breve{\rho}_R \operatorname{sgn}(\mu_R - \sigma) \Big). \quad (34)$$

The sums in the rational polynomial $G_i(v)$ contain different $\breve{\rho}$ constants but the same set of $\operatorname{sgn}(\mu_R - \sigma)$. Hence, it is written as $\breve{G}$ for shorthand.

Then, let $\sigma = v_2$ in order to write the one-dimensional integral in (32). The solution is given as a sum of $L_k^i + 1$ terms as

$$I_2(v_1) = \sum_{m=1}^{L_k^i+1} \exp\Big(-\sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \rho_R |\mu_R - \mu_m| + j\xi_k^i \mu_m \Big) \times \frac{\breve{G}_i\big(+\breve{\rho}_m + \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \breve{\rho}_R \operatorname{sgn}(\mu_R - \mu_m)\big)}{j\xi_k^i + \rho_m + \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \rho_R \operatorname{sgn}(\mu_R - \mu_m)} - \frac{\breve{G}_i\big(-\breve{\rho}_m + \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \breve{\rho}_R \operatorname{sgn}(\mu_R - \mu_m)\big)}{j\xi_k^i - \rho_m + \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \rho_R \operatorname{sgn}(\mu_R - \mu_m)}. \quad (35)$$

This complicated looking (35) simplifies readily into a simple, double-sided scalar integral over v_1,

$$I_2(v_1) = \sum_{m=1}^{L_k^i+1} \{a_{i,m} + j\, d_{i,m} \operatorname{sgn}(v_1)\} \cdot e^{-D_m \cdot |v_1| + j\xi_k^i v_1}. \quad (36)$$

This simplification is based on algebraic relations used in the estimator's measurement update process in [7], as well as constants $D_{i,m}$, $\bar{D}_{i,m}$, and $\breve{D}_{i,m}$ defined as

$$D_{i,m} \cdot |v_1| = \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \rho_R |\mu_R - \mu_m| \quad (37a)$$

$$\bar{D}_{i,m} \cdot \operatorname{sgn}(v_1) = \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \rho_R \operatorname{sgn}(\mu_R - \mu_m) \quad (37b)$$

$$\breve{D}_{i,m} \cdot \operatorname{sgn}(v_1) = \sum_{\substack{R=1 \\ R \neq m}}^{L_k^i+1} \breve{\rho}_R \operatorname{sgn}(\mu_R - \mu_m) \quad (37c)$$

$$\hat{\xi}_k^i \cdot v_1 = \xi_k^i \cdot \mu_m. \quad (37d)$$

Denote the outer integral with respect to v_1 in (30) by

$$I_1 = \sum_{m=1}^{L_k^i+1} \int_{-\infty}^{\infty} \{a_{i,m} + j\, d_{i,m} \operatorname{sgn}(v_1)\} \times \exp\Big(-D_m \cdot |v_1| + j\hat{\xi}_k^i v_1 \Big) \exp(-\eta_1 |v_1| + j\bar{x}_1 v_1) \times \exp\Big(j \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{R,i}} \Big) dv_1$$

$$= \sum_{m=1}^{L_k^i+1} \int_{-\infty}^{\infty} \{a_{i,m} + j\, d_{i,m} \operatorname{sgn}(v_1)\} \times \exp\Big(-D_m + \eta_1 \Big) |v_1| + j\Big(\hat{\xi}_k^i + \bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big) v_1 \Big) dv_1. \quad (38)$$

This integral has a form identical to the measurement update process for a scalar system [4]. Its solution is given by

$$I_1 = \sum_{m=1}^{L_k^i+1} \frac{a_{i,m} - j\, d_{i,m}}{j\Big(\hat{\xi}_k^i + \bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big) + D_m + \eta_1} - \frac{a_{i,m} + j\, d_{i,m}}{j\Big(\hat{\xi}_k^i + \bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big) - D_m + \eta_1}$$

$$= \sum_{m=1}^{L_k^i+1} 2 \cdot \frac{a_{i,m} D_m + \eta_1 - d_{i,m} \Big(\bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big)}{\Big(\hat{\xi}_k^i + \bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big)^2 + D_m + \eta_1^2}. \quad (39)$$

Finally, the conditional performance index in (30) is given by

$$\tilde{J}_{Z_k} = \frac{1}{2\pi^2} \int_{-\infty}^{\infty} \frac{\zeta_{p-1}/\pi}{u^2(i) + \zeta_i^2} \times \sum_{i=1}^{N} \sum_{m=1}^{L_k^i+1} \frac{a_{i,m} D_m + \eta_1 - d_{i,m} \Big(\bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big)}{\Big(\hat{\xi}_k^i + \bar{x}_1 + \sum_{R=1}^{2} Z_{k,i}^R B_{k,1}^{M_{k,i}^R} \Big)^2 + D_m + \eta_1^2} \Big|_{i=k}. \quad (40)$$

This closed form conditional performance index is non-convex and depends on the control input sequence $\{u(k), \ldots, u(p-1)\}$ in a complex way; specifically, the

parameters $a_{i,m}$, $d_{i,m}$, and $\bar{x}_1$ depend on the control sequence. Thus, we maximize (40) numerically using the accelerated gradient search method [14]. The optimization is done in two steps: first, the global optimum of the double sum term in (40) without the control weighting terms is optimized with respect to $\begin{bmatrix}\bar{x}_1 & \bar{x}_2\end{bmatrix}^T = \bar{x}(p)$; then, that final state is used to generate a control sequence as an initial guess for the second, local accelerated gradient search optimization step.

VI. Numerical Examples

Here, we present two sets of examples, the first of which shows the optimal control $u(0)^*$ versus the measurement $z(0)$ for the first measurement update only, and the second set shows two multi-step examples. All of these examples use a two-step horizon, i.e. $m = 2$, so that there exists a control sequence that can drive our two-state system to the origin over this horizon length. However, as we are using model predictive control, only the first control input of this sequence is applied at that time step.

All of our examples compare our Cauchy optimal model predictive controller with a similar LEG model predictive optimal controller. The LEG estimator assumes that the stochastic inputs are described by the Gaussian pdfs that are closest, in a least squared sense, to the given Cauchy pdfs; and the LEG controller assumes that its objective functions of the state and control resemble scaled Gaussian pdfs that are closest, in the least-squared sense, to the scaled Cauchy pdfs in (10). The LEG controllers' responses are shown in dashed lines in the figures.

The first set of examples are shown in Fig. 1. These figures show the applied optimal control input at the first time step, $k = 0$, given the first measurement. In the two cases presented, all the systems parameters are the same, except in Fig. 1(a) $\gamma > \alpha_1 = \alpha_2$ (i.e. more measurement than state uncertainty), and in Fig. 1(b) $\alpha_1 = \alpha_2 > \gamma$ (i.e. more state than measurement uncertainty).

The example in Fig. 1(a) shows that the Cauchy controller reduces its control effort to zero as the measurement deviations become large. This is in contrast to the LEG controller, which is linear and thus responds strongly to large measurement deviations. This behavior in the Cauchy controller occurs when the measurement uncertainty is larger than the state uncertainty. In the opposite case shown in Fig. 1(b), the measurement has less uncertainty than the state. In this case, the Cauchy controller's response closely matches that of the LEG in a neighborhood of the origin, and in fact responds even more strongly than the LEG for large measurement deviations.

The three different curves in both of these figures represent the response for three different control weights: no control weight, $\zeta_0 = 10$, and $\zeta_0 = 5$. As expected, heavier control weights (i.e. smaller ζ) reduce the control effort, but even without any control weighting, the response in Fig. 1(a) goes to zero for large measurement deviations. The fact that this behaviour is seen when there is no control weighting implies that the attenuation of the control signal for large measurement deviations is due to the cpdf and not the objective function.

The complexity of evaluating the cost grows as the number of terms increases across time steps, as indicated in (40). For implementable control, this growth needs to be arrested. The full information characteristic function of the ucpdf (25) is approximated by a characteristic function of a ucpdf conditioned on a fixed sliding window of the most recent measurements, taken here to have length eight. The relative error in the approximation appears to be 10^{-6} or smaller.

The second set of examples are shown in Fig. 2. All of the plots show the state, control, and noise histories for the given simulations. The difference between Fig. 2(a) and Fig. 2(b) is that the process noise β and the measurement noise γ parameter values are interchanged.

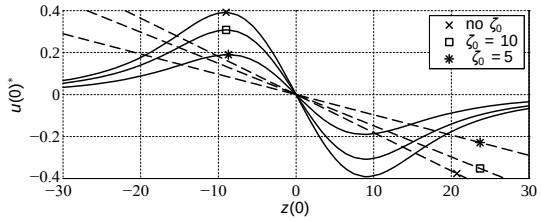
It is interesting to compare Fig. 2(a) with Fig. 2(b) in light of Fig. 1. In Fig. 2(a), there is more uncertainty in the state process noise than in the measurement noise. When large measurement deviations occur (such as at $k = 52$), the Cauchy controller's effort is very small. In contrast, the LEG controller responds with a large control effort that drives the states from their regulated state of zero. However, when large process noise inputs occur, the state deviates and the Cauchy controller applies a larger control effort than the LEG, thus regulating the state more effectively. This suggests that when the measurement noise density parameter dominates the process noise density parameter in constructing the Cauchy controller, the effect of measurement outliers is mitigated, while still responding to state deviations due to process noise.

On the other hand, in Fig. 2(b) there is more uncertainty in the state than in the measurement, and the Cauchy controller behaves very much like the linear LEG controller. The state trajectories and control inputs of the Cauchy and LEG controllers appear equal, but actually their differences are much smaller than the scale of the axis and cannot be seen. This suggests that, when the stochastic parameters allow it, the Cauchy controller follows the measurement more. It responds in a more linear fashion to the measurements, imitating the performance of the LEG controller in that setting.

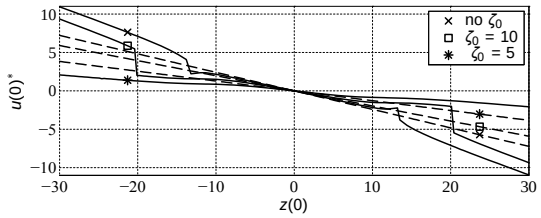
This behavior is seen again when both controllers face Gaussian noises, as in Figure 2(c). Here, the Cauchy controller closely follows both the control and state trajectories of the LEG, which is the true optimal solution. Hence, the Cauchy is robust under non-impulsive noise environments, as it closely approximates the true optimal solution given by the LEG.

VII. Conclusions

An optimal stochastic controller was derived for vector-state, linear, discrete-time systems with additive process and measurement Cauchy distributed noises. Since the Cauchy distribution has an undefined mean and an infinite second moment, we cannot use standard objective function, e.g., the expected value of a quadratic function of state and control variables. Therefore, a new and computable objective function was defined. Opposed to previous work, the characteristic function of the cpdf of the state given the measurement history and the Parseval's equation are used to express the



(a) optimal control vs the measurement: $\alpha_i = 0.1$, $\gamma = 0.5$.



(b) optimal control vs the measurement: $\alpha_i = 0.5$, $\gamma = 0.1$.

Fig. 1. Parameters used are: $\eta = [1\ 1]$, $\beta = 0.02$, $\Gamma^{\mathrm{T}} = [1\ 1]$, $\Lambda^{\mathrm{T}} = [1\ 1]$, and the eigenvalues of Φ are $0.2455 \pm j0.1523$.

conditional performance index in a closed form. This closed-form conditional performance index is optimized numerically using an accelerated gradient search. Examples are presented that show how our vector state Cauchy controller compares against an equivalent LEG controller, demonstrating the Cauchy controller's performance and improved robustness over its Gaussian counterpart.

REFERENCES

[1] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution," *IEEE Communications Letters*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[2] P. Reeves, "A non-gaussian turbulence simulation," Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-69-67, 1969.

[3] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall, 1994.

[4] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1329–1342, 2010.

[5] J. L. Speyer, M. Idan, and J. Fernández, "Multi-Step Prediction Optimal Control for a Scalar Linear System with Additive Cauchy Noise," in *IEEE Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

[6] M. Idan and J. L. Speyer, "State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach," *SIAM J. Control Optim.*, vol. 50, no. 4, pp. 1971–1994, 2012.

[7] J. L. Speyer, M. Idan, and J. Fernández, "The Two-State Estimator for Linear System with Additive Measurement and Process Cauchy Noise," in *IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2012.

[8] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. New York: Springer-Verlag, 1975, ch. 5 & 6.

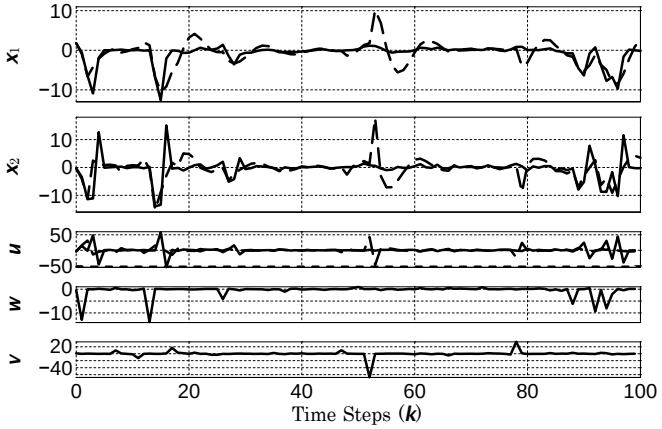
[9] J. L. Speyer, M. Idan, and J. Fernández, "A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise," *Submitted to Automatica*, 2013.

[10] M. Morari and E. Zafiriou, *Robust Process Control*. New Jersey: Prentice-Hall, 1989.

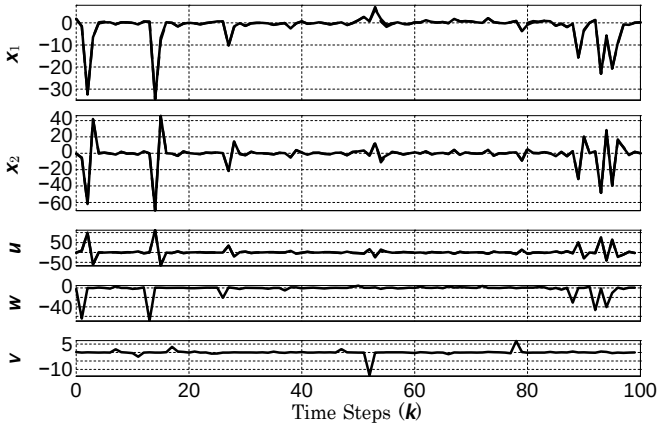
[11] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

[12] M. Idan and J. L. Speyer, "Multivariate cauchy estimator with scalar measurement and process noises," *Submitted to SIAM*, 2013.

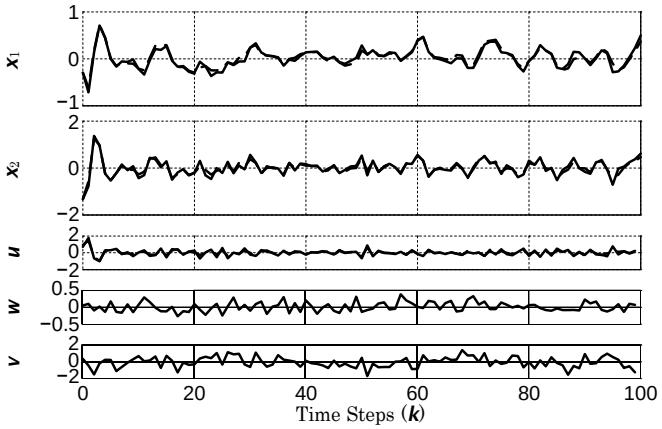
[13] ——, "Multivariate Cauchy Estimator with Scalar Measurement and Process Noises," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Florence, Italy, December 2013.



(a) $\gamma = 0.5$, $\beta = 0.1$, Cauchy noises.



(b) $\gamma = 0.1$, $\beta = 0.5$, Cauchy noises.



(c) $\gamma = 0.5$, $\beta = 0.1$, Gaussian noises.

Fig. 2. Cauchy and LEG controllers' performance, solid lines are Cauchy and dashed lines are LEG. Simulation parameters are: $\eta = [1\ 1]$, $\alpha = [0.8\ 0.8]$, $\mathrm{H} = [1\ 1]$, $\Gamma^{\mathrm{T}} = [0.5\ 1]$, $\Lambda^{\mathrm{T}} = [0.5\ 1]$, $m = 2$, and the eigenvalues of Φ are $0.8 \pm j0.55$. There is no control weighting.

[14] R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," *The Computer Journal*, vol. 6, no. 2, pp. 163–168, 1963.

F Appendix

Javier H. Fernández, Jason L. Speyer, and Moshe Idan
Linear Dynamic Systems with Additive Cauchy Noises Part 2: Stochastic Model Predictive Controls
Submitted to IEEE Transactions on Automatic Control

Linear Dynamic Systems with Additive Cauchy Noises
Part 2: Stochastic Model Predictive Control

Javier H. Fernández *Student Member, IEEE*, Jason L. Speyer *Life Fellow, IEEE*, and Moshe Idan

***Abstract*—An optimal predictive controller for linear, vector-state dynamic systems driven by Cauchy measurement and process noises is developed. For the vector-state system, the probability density function (pdf) of the state conditioned on the measurement history cannot be generated. However, the characteristic function (CF) of this pdf can be expressed in an analytic form. Consequently, the performance index used for the controller design is evaluated in the spectral domain using this CF. By taking the conditional expectation of an objective function that is a product of functions resembling Cauchy pdfs, the conditional performance index is obtained analytically in closed form by using Parseval's identity and integrating over the spectral vector. This forms a deterministic, non-convex function of the control signal and the measurement history that must be optimized numerically at each time step. A two-state example is used to expose the interesting robustness characteristics of the proposed controller.**

I. Introduction

Control of dynamic systems in real world applications, from engineering and science to economics and finance, frequently involves handling uncertain, stochastic inputs. These uncertainties affect both the actual state of the system as well as the measurements that the controller depends on. When designing controllers for stochastic systems, it is often assumed that the uncertainties are described by Gaussian probability density functions (pdf), due to the efficiency with which modern methods handle them. For linear systems, algorithms like the linear quadratic Gaussian (LQG) and the linear exponential Gaussian (LEG) assume linear dynamics and additive process and measurement noises described by the Gaussian pdf [1].

In many applications the underlying random processes have an impulsive character producing deviations of high amplitude and small duration much more often than the Gaussian assumption permits [2]. Examples of such processes include radar and sonar noise [3] and disturbances due to air turbulence [4]. Another application is adversarial missile guidance, where the target is intelligent and desires to evade missile. In missile guidance, uncertainties are usually assumed to be Gaussian. However, optimal target evasion maneuvers involve high acceleration rate maneuvers, i.e. a high amplitude and small duration input [5]. Hence, a control algorithm assuming that the target is driven by a light tailed distribution would not capture this behavior. The Gaussian distribution function has very light tails, so that large deviations are extremely unlikely. Therefore, the LQG (H_2) and LEG (H_∞) algorithms do not perform well in the presence of heavy-tailed or impulsive uncertainties.

Impulsive uncertainties, like those mentioned in the examples above, are better described by heavy-tailed distributions, such as the symmetric alpha-stable ($S\alpha S$) distributions [6]. These distributions are described not by their probability density functions (pdfs), but by their characteristic functions (CFs). They are of the form $\varphi(v) = e^{-\sigma^\alpha |v|^\alpha + j\mu v}$, where σ is the scaling parameter, μ is the median, v is the spectral variable, and the characteristic exponent α determines the type of distribution: $\alpha = 2$ implies the Gaussian distribution, and $\alpha = 1$ implies the Cauchy distribution.

Estimation assuming Cauchy distributed noises has shown improved performance over Gaussian estimators when faced with impulsive noises. For estimating the direction of arrival of a signal to a sensor array in [7], maximum likelihood estimators designed assuming Cauchy distributed noises were shown to exhibit performance very close to the Cramèr-Rao Bound against $S\alpha S$ noises with characteristic exponents $1 \le \alpha \le 2$. Similar performance was observed in various applications, including processing data in a multi-user communication network [8] and radar glint [9]. However, these studies are only for estimation of the signal, or signal parameters, and do not attempt to control their respective systems. This paper derives a framework for control of multivariate linear systems driven by Cauchy distributed measurement and process noises.

Algorithms for optimal estimation and control of scalar linear systems driven by Cauchy distributed process and measurement noises have been developed in [10, 11]. There, the conditional performance index for model predictive control is determined by taking the conditional expectation of the objective function using the probability density given the measurement history as presented in [11, 12]. A dynamic programming algorithm is also developed in [12]. It is shown that the solution to the dynamic programming recursion is intractable because of the need to average over future measurements in determining the optimal return function. This cannot be done in closed form due to the complex dependency of the optimal return function on the measurement history. Hence, the dynamic programming solution was approximated using the model predictive control method.

In this paper, the Cauchy optimal control algorithm for scalar systems [12] is extended to systems with a vector state. For the vector case, the conditional pdf (cpdf) given the measurement history is not available. However, the CF of the cpdf can be recursively propagated for such vector state systems [13–17]. In particular, [16, 17] presents an efficient algorithm for the two-state Cauchy estimator. Although this control methodology can be applied to general vector-state systems in [13, 14], the closed form expression for the conditional performance index used in the control problem will be based on the structure presented in [16, 17].

The significant contribution of this paper is evaluating *in closed form* the conditional performance index using the cpdf's CF instead of the cpdf as in [12]. Preliminary results can be found in [18]. The objective function is cast as a product of functions resembling Cauchy pdfs, which are easily transformed into functions of the spectral variables. Consequently, the conditional performance index, found in a closed form, is a deterministic function of the control and measurement histories. Due to its complexity, the optimal control signal is determined by numerically optimizing this non-convex conditional performance index in a model predictive control setting.

The remainder of the paper is structured as follows. The controlled system model is presented in Section II. An appropriate, computable performance index for this problem is presented in Section III and

This work was partially supported by Air Force Office of Scientific Research, Award No. FA9550-09-1-0374, and by the United States - Israel Binational Science Foundation, Grant 2008040.

J. H. Fernández is with the department of Mechanical and Aerospace Engineering, University of California, Los Angeles.

Email: jhf@seas.ucla.edu.

J. L. Speyer is with the department of Mechanical and Aerospace Engineering, University of California, Los Angeles.

Email: speyer@seas.ucla.edu.

M. Idan is with the Faculty of Aerospace Engineering, Technion, Haifa, Israel. Email: moshe.idan@technion.ac.il.

subsequently transformed from the state variable to the spectral variable form. In Section IV the spectral integrations required to determine the conditional performance index are reduced to an integral formula that can be evaluated in closed form. Section V addresses a special case of systems with two states, using the efficient algorithm presented in [17]. Here, using an alternative, simplified form of the two state cpdf's CF, the conditional performance is determined in closed form. In Section VI numerical examples are given, demonstrating the performance of the Cauchy controller under both Cauchy and Gaussian noise conditions. Conclusions are given in Section VII.

II. Description of the Model

This paper deals with a discrete time, linear system described by

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Lambda u(k) + \Gamma w(k) \\ z(k) &= \mathrm{H} x(k) + v(k) \end{aligned} \tag{1}$$

where $x(k) \in \mathrm{R}^n$ is the state vector, $u(k)$ is a scalar input, $z(k)$ is a scalar measurement, and $w(k)$ and $v(k)$ are scalar independent Cauchy distributed process and measurement noise inputs with medians at zero and scaling parameters of β and γ, respectively, so that their pdfs are given by

$$f_{W}\ w(k) = \frac{\beta/\pi}{w^2(k)+\beta^2},\quad f_{V}\ v(k) = \frac{\gamma/\pi}{v^2(k)+\gamma^2}. \tag{2}$$

The CFs of these pdfs are

$$\varphi_{W}(\sigma) = e^{-\beta|\sigma|},\quad \varphi_{V}(\sigma) = e^{-\gamma|\sigma|}, \tag{3}$$

where σ is the scalar spectral variable.

The initial conditions are assumed to be independent Cauchy distributed random variables with the pdfs

$$f_{X_1}(x(1)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{(x_i(1)-\bar{x}_i(1))^2+\alpha_i^2}. \tag{4a}$$

Its CF is given by

$$\varphi_{X_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i|v_i| + j\bar{x}_i(1)v_i}, \tag{4b}$$

where v_i is an element of $v \in \mathrm{R}^n$.

The stochastic system (1) can be decomposed into two systems, one driven by $u(k)$ and one by $w(k)$, by exploiting the linearity of the system. Let $\bar{x}(k)$ and $\bar{z}(k)$ be the part of the system driven by the input $u(k)$ only, and $\tilde{x}(k)$ and $\tilde{z}(k)$ be the part of the system driven by the process noise $w(k)$ only and contains all the underlying random variables. Then,

$$x(k) = \bar{x}(k) + \tilde{x}(k) \tag{5a}$$

$$z(k) = \bar{z}(k) + \tilde{z}(k). \tag{5b}$$

The controlled part of the system is described by

$$\bar{x}(k+1) = \Phi\bar{x}(k) + \Lambda u(k) \tag{6a}$$

$$\bar{z}(k) = \mathrm{H}\bar{x}(k) \tag{6b}$$

with initial condition $\bar{x}(1)$, i.e., the median of (4a). The process noise driven part is given by

$$\tilde{x}(k+1) = \Phi\tilde{x}(k) + \Gamma w(k) \tag{7a}$$

$$\tilde{z}(k) = \mathrm{H}\tilde{x}(k) + v(k). \tag{7b}$$

The process and measurements noise pdfs were defined in (2), while the initial condition of this stochastic model is Cauchy distributed with a pdf given by

$$f_{\tilde{X}_1}(\tilde{x}(1)) = \prod_{i=1}^{n} \frac{\alpha_i/\pi}{\tilde{x}_i^2(1)+\alpha_i^2}. \tag{8a}$$

Its CF is

$$\varphi_{\tilde{X}_1}(v) = \prod_{i=1}^{n} e^{-\alpha_i|v_i|}. \tag{8b}$$

The above decomposition will be used to derive the Cauchy controller.

III. Derivation of the Cost Using Characteristic Functions

Our proposed controller is an m-step model predictive controller [19] that uses current and past measurements, and averages over future process noise. At each time step, the conditional performance index is computed. Since the performance index was found to be a non-convex function of the control sequence, it is maximized numerically. Once the optimal control sequence of length m is computed, only the first control in that sequence is applied. At the next step, a new measurement is taken and the process is repeated, producing a new optimal control sequence and applying only the first one. In this paper, we study the optimal stochastic state regulation problem, noting that the tracking problem can be handled in a similar fashion. Our regulation problem will have a finite horizon of length m such that the terminal state occurs at time-step $p = k + m$.

Let the state, measurement, and control histories used in the control problem formulation be defined as

$$X_{k+1}^{p} := \{x(k+1), \ldots, x(p)\}, \tag{9a}$$

$$Z_k := \{\tilde{z}(1), \ldots, \tilde{z}(k)\}, \tag{9b}$$

$$U_k^{p-1} := \{u(k), \ldots, u(p-1)\},\ U_k^{p-1} \in F, \tag{9c}$$

where F is the class of piecewise continuous functions adapted to the σ-algebra σ_k generated by the measurement history, i.e. the control is a random variable that is measurable with respect to events in σ_k [20]. Moreover, in [12] it is proven that u_k is adapted to the σ-algebra $\tilde{\sigma}_k$ generated by $\tilde{Z}_k$, which means that the control is measurable on events generated by $\tilde{Z}_k$ only.

Similar to the scalar control problem presented in [11, 12], the control objective function is chosen as a product of functions resembling Cauchy pdfs, given by

$$\psi\left(X_{k+1}^{p}, U_{k}^{p-1}\right) = \prod_{i=k}^{p-1} \frac{\zeta_i/\pi}{u^2(i)+\zeta_i^2} \cdot \prod_{r=1}^{n} \frac{\eta_{i+1,r}/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2}. \tag{10}$$

Then, the the performance index conditioned on the current measurement history and averaged over future process noises is given by

$$\begin{aligned} J_{k,p}^{*} &= \max_{U_k^{p-1}\in F} E\left[\psi\left(X_{k+1}^{p}, U_{k}^{p-1}\right) \,\middle|\, \tilde{Z}_k\right] \\ &= \max_{U_k^{p-1}\in F} E\left[E\left[\psi\left(X_{k+1}^{p}, U_{k}^{p-1}\right) \,\middle|\, Z_k\right] \,\middle|\, \tilde{Z}_k\right], \\ &= E\left[\max_{U_k^{p-1}\in F} E\left[\psi\left(X_{k+1}^{p}, U_{k}^{p-1}\right) \,\middle|\, Z_k\right] \,\middle|\, \tilde{Z}_k\right] = E\left[J_{k}^{*} \,\middle|\, \tilde{Z}_k\right] \end{aligned} \tag{11}$$

where the interchange of the maximum and expectation operations is due to the fundamental theorem [1].

$_{i=1}$ $_{i}(1) + \alpha_{i}$

We are now concerned with determining an analytic form for the

conditional performance index J_{Z_k}, where using (10) it becomes

$$\begin{aligned} J_{Z_k} &= E_{X_{k+1},U_k^{p-1}|Z_k}\Big[\psi(p)\prod_{i=k}^{p-1}\frac{\eta_{i+1,r}^2/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2}\cdot\frac{\zeta_i^2/\pi}{u^2(i)+\zeta_i^2}\Big] \\ &= \int_{-\infty}^{\infty}\prod_{i=k}^{p-1}\prod_{r=1}^{n}\frac{\eta_{i+1,r}^2/\pi}{x_r^2(i+1)+\eta_{i+1,r}^2}\cdot\frac{\zeta_i^2/\pi}{u^2(i)+\zeta_i^2} \\ &\quad\times f_W(\tilde{x}(p)|\tilde{x}(p-1))\cdots f_W(\tilde{x}(k+2)|\tilde{x}(k+1)) \\ &\quad\times f_{\tilde{X}_{k+1}|Z_k}(\tilde{x}(k+1)|Z_k)d\tilde{x}_1(k+1)\ldots d\tilde{x}_n(k+1) \\ &\quad\times d\tilde{x}_1(k+2)\ldots d\tilde{x}_n(k+2)\ldots d\tilde{x}_1(p)\ldots d\tilde{x}_n(p). \end{aligned} \tag{12}$$

For presentation simplicity, in this derivation we will only consider weighting on the terminal state, $x(p)$, and on the np scalar control inputs. The control weighting functions $M^{\prod_{i=k}^{p-1}\zeta_i^2/\pi}$ can come out of the integral in (12). Then, the product over i inside the integrand has only the term for $i = p-1$. For notational convenience we can drop the time-step index in this integral and write it over $\{\tilde{x}_1,\ldots,\tilde{x}_n\}$ as

$$\begin{aligned} J_{Z_k} &= M^{U}\int_{-\infty}^{\infty}\prod_{r=1}^{n}\frac{\eta_{p,r}^2/\pi}{\tilde{x}_r+\bar{x}_r)^2+\eta_{p,r}^2}\,f_{X_p|Z_k}(\tilde{x}|Z_k)d\tilde{x}_1\ldots d\tilde{x}_n \\ &= M^{U}\int_{-\infty}^{\infty}R_x(\tilde{x}+\bar{x})\cdot f_{\tilde{X}_{p|Z_k}}(\tilde{x}|Z_k)d\tilde{x}_1\ldots d\tilde{x}_n. \end{aligned} \tag{13}$$

The cpdf $f_{\tilde{X}_p|Z_k}(\tilde{x}|Z_k)$ can be evaluated in closed form for scalar systems [10]. However, for vector state systems only the CF of the cpdf in (13), $\varphi_{p|k}(v)$, can be evaluated in closed form [13]. Therefore, when computing the conditional performance index we need to express the integral using the spectral variable v instead of the pdf variable $\tilde{x}$. The CF is the inverse Fourier transform of a pdf, and therefore CFs retain all the properties of Fourier transforms. Hence, we use the same transform on the objective function R_x in (13), and then apply Parseval's identity. The inverse Fourier transform of R_x in (13) is denoted L_x and given by

$$L_x(v) = \int_{-\infty}^{+\infty}R_x(\tilde{x}+\bar{x})e^{jv^{\mathrm{T}}\tilde{x}}d\tilde{x} = \prod_{r=1}^{n}e^{-\eta_{p,r}|v_r|-j\bar{x}_r}. \tag{14}$$

Using these definitions, we can apply Parseval's identity over each variable in (13) to express the conditional performance index as an integral over the spectral variable v,

$$\begin{aligned} J_{Z_k} &= \frac{M^U}{(2\pi)^n}\int_{-\infty}^{\infty}L_x^*(v)\cdot\varphi_{p|k}(v)\,dv_1\ldots dv_n \\ &= \frac{M^U}{(2\pi)^n}\int_{-\infty}^{\infty}\Big(\prod_{r=1}^{n}e^{-\eta_{p,r}|v_r|+j\bar{x}_r(p)v_r}\Big)\varphi_{p|k}(v)\,dv_1\ldots dv_n, \end{aligned} \tag{15}$$

IV. The Conditional Performance Index

Consider the integral over v_n in (15),

$$\begin{aligned} I_n &= \int_{-\infty}^{\infty}\Big(\prod_{r=1}^{n}e^{-\eta_{p,r}|(E_r,v)|}\Big)e^{j\langle\bar{x}(p),v\rangle}\varphi_{p|k}(v)dv_n \\ &= \int_{-\infty}^{\infty}e^{-\sum_{r=1}^{n}\eta_{p,r}|(E_r,v)|+j\langle\bar{x}(p),v\rangle}\varphi_{p|k}(v)dv_n, \end{aligned} \tag{16}$$

where E_r is the r^{th} column of the n-dimensional identity matrix. The cpdf for the state $\tilde{x}(k)$ is denoted as $f_{\tilde{X}_k|\tilde{Z}_k}$. The unnormalized cpdf (ucpdf) is denoted as $f_{\tilde{X}_k|\tilde{Z}_k}=f_{\tilde{X}_k|\tilde{Z}_k}\cdot f_{\tilde{Z}_k}$, where $f_{\tilde{Z}_k}$ is the pdf of the measurement history and has a known value. In [13,14], the CF of the ucpdf $\varphi_{p|k}(v)$ is recursively propagated; the CF of the normalized cpdf is $\varphi_{k|k}(v)=\varphi_{k|k}(v)/f_{\tilde{Z}_k}$, where $f_{\tilde{Z}_k}=\varphi_{\tilde{X}_k|\tilde{Z}_k}(v)|_{v=0}$. From [13, 14] the form of the CF of the ucpdf at time k is shown to be

$$\varphi_{k|k} = \sum_{i=1}^{n_t^{k|k}}g_i^{k|k}(v_{gi}(v))e^{y_{ei}^{k|k}(v)}, \tag{17a}$$

where

$$v_{gi}^{k|k}(v) = \sum_{\ell=1}^{n_{ei}^{k|k}}q_{i,\ell}^{k|k}\operatorname{sgn}\big(\langle a_{i\ell}^{k|k},v\rangle\big)\in\mathbb{R}^k, \tag{17b}$$

$$v_{ei}^{k|k}(v) = -\sum_{\ell=1}^{n_{ei}^{k|k}}p_{i,\ell}^{k|k}\big|\langle a_{i\ell}^{k|k},v\rangle\big| + j\langle b_i^{k|k},v\rangle, \tag{17c}$$

and the parameters $n_t^{k|k}$, $n_{e,i}^{k|k}$, $q_{i\ell}^{k|k}$, $p_{i\ell}^{k|k}$, $a_{i\ell}^{k|k}$, $b_i^{k|k}$ are generated sequentially from $k=1$.

For the MPC algorithm, the CF of the ucpdf is to be propagated through the stochastic dynamics (7) to time $k+m=p$ using the propagation formula given in [13, 17]. The CF of the m-step propagated cpdf is denoted $\bar{\varphi}_{p|k}(v)$ and given by

$$\begin{aligned} \bar{\varphi}_{p|k}(v) &= \varphi_{\tilde{X}_k|\tilde{Z}_k}(\Phi^{m\mathrm{T}}v)\varphi_{n_t^{k|k}}((\Phi^{m-1}\Gamma)^{\mathrm{T}}v)\cdots\varphi_{n_t^{k|k}}((\Phi\Gamma)^{\mathrm{T}}v)\varphi_{n_t^{k|k}}(\Gamma^{\mathrm{T}}v) \\ &= \prod_{i=1}^{n_t^{k|k}}g_i^{k|k}(v_{gi}(\Phi^{m\mathrm{T}}v))e^{y_{ei}^{k|k}(\Phi^{m\mathrm{T}}v)} \\ &\quad\times\exp\Big(-\beta\big|\langle\Phi^{m-1}\Gamma,v\rangle\big|-\cdots-\beta\big|\langle\Phi\Gamma,v\rangle\big|-\beta\big|\langle\Gamma,v\rangle\big|\Big). \end{aligned} \tag{18}$$

In (18) we effectively add m terms to the sum in $y_{ei}^{k|k}(v)$ of (17c). By combining the exponent in (16) with that in (18), the combined exponent in the integrand of (16) has in total $n_{ei}^{k|k}+m+n$ real terms, while the imaginary part is composed of two components. Define the following terms

$$\begin{aligned} &\bar{p}_{i\ell}=p_{i\ell}, && \bar{a}_{i\ell}^{k|k}=\Phi^m a_{i\ell}^{k|k} && \text{for } R=1,\ldots,n_{ei}^{k|k} \\ &\bar{p}_{i\ell}=\beta, && \bar{a}_{i\ell}^{k|k}=\Phi^t\Gamma && \text{for } \begin{matrix}R=1,\ldots,n_{ei}^{k|k},\\ t=R-(n_{ei}^{k|k}+1)\end{matrix} \\ &\bar{p}_{i\ell}=\eta_{p\ell}, && \bar{a}_{i\ell}^{k|k}=E_r && \text{for } \begin{matrix}R=1,\ldots,n_{ei}^{k|k},\\ r=R-(n_{ei}^{k|k}+m)\end{matrix} \end{aligned} \tag{19a}$$

and

$$\begin{aligned} &\bar{b}_i^{k|k}=b_i^{k|k}+\bar{x}_p, && \bar{n}_{ei}^{k|k}=n_{ei}^{k|k}+m+n_{k|k} \\ &q_{t\ell}=0 && \text{for } R=n_{ei}^{k|k}+1,\ldots,\bar{n}_{ei} \end{aligned} \tag{19b}$$

Using these definitions, the integrand in (16) becomes

$$\prod_{r=1}^{n}e^{-\eta_{p,r}|(E_r,v)|+j\langle\bar{x}(p),v\rangle}\varphi_{p|k}(v) = \prod_{i=1}^{n_t^{k|k}}g_i^{k|k}(v_{gi}(v))e^{y_{ei}^{k|k}(v)}, \tag{20}$$

where L_x^* is the complex conjugate of L_x. The next section shows how to evaluate these n nested integrals sequentially in closed form.

$$\psi(v) = \prod_{i=1}^{n} \psi_i(v) = \prod_{i=1}^{n} g_i^{k|k}\left(\bar{y}_{gi}^{k|k}(v)\right) \cdot e^{\bar{y}_{ei}^{k|k}(v)}, \tag{20a}$$

where

$$\bar{v}_{gi}^{k|k}(v) = \sum_{\ell=1}^{\bar{r}_{gi}^{k|k}} q_{i\ell}^{k|k} \operatorname{sgn}\left(\left\langle \bar{a}_{i\ell}^{k|k}, v \right\rangle\right), \tag{20b}$$

$$\bar{y}_{ei}^{k|k}(v) = \sum_{\ell=1}^{\bar{r}_{ei}^{k|k}} \bar{p}_{i\ell}^{k|k} \left| \left\langle \bar{a}_{i\ell}^{k|k}, v \right\rangle + j\, b_{i\ell}^{k|k} \right|. \tag{20c}$$

The integration in (16) is performed for each element of v in turn. Beginning with v_n, decompose $v = \begin{bmatrix} \hat{v} & v_n \end{bmatrix}^{\mathrm{T}}$, where $v \in \mathbb{R}^{n-1}$. Then,

$$I_n = \int_{-\infty}^{\infty} \psi(v)\,dv = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{\bar{n}_{k|k}} \psi_i(v)\,dv$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{i=1}^{\bar{n}_{k|k}} \psi_i(v_n, \hat{v}) \right] dv_n\, d\hat{v}. \tag{21}$$

The objective is to reduce the inner integral in (21) to a form that is obtained in closed form using the integral formula developed in [13, 14]. First, since $\bar{a}_{i\ell}^{k|k}$ multiplies v in the sign function in (20b) and the absolute value function in (20c), they are decomposed as $\bar{a}_{i\ell}^{k|k} = \begin{bmatrix} \hat{a}_{i\ell}^{k|k} & \tilde{a}_{i\ell}^{k|k} \end{bmatrix}^{\mathrm{T}}$, where $\tilde{a}_{i\ell}^{k|k}$ is a scalar and $\hat{a}_{i\ell}^{k|k} \in \mathbb{R}^{n-1}$. Therefore, the inner products in (20b) and (20c) become

$$\left\langle \bar{a}_{i\ell}^{k|k}, v \right\rangle = \left\langle \hat{a}_{i\ell}^{k|k}, \hat{v} \right\rangle - \tilde{a}_{i\ell}^{k|k} v_n. \tag{22}$$

In order to rewrite (22) in a form consistent with the integral formula in [13, 14], $-\tilde{a}_{i\ell}^{k|k}$ is factored out of the second term. If $\tilde{a}_{i\ell}^{k|k} = 0$, then the term $e^{\tilde{a}_{i\ell}^{k|k} v_n}$ loses dependence on v_n and it is removed from the inner integral in (21). Therefore, only $\tilde{a}_{i\ell}^{k|k} \neq 0$ needs to be considered. Let (22) be rewritten as

$$\left\langle \bar{a}_{i\ell}^{k|k}, v \right\rangle = \tilde{a}_{i\ell}^{k|k} \operatorname{sgn}\left(-\tilde{a}_{i\ell}^{k|k}\right) \left(\mu_{i\ell}^{k|k} - v_n \right), \tag{23a}$$

where $\mu_{i\ell}^{k|k} = \left\langle -\bar{a}_{i\ell}^{k|k}, \hat{v} \right\rangle$. Therefore, the elements in (20b) and (20c) are

$$q_{i\ell}^{k|k} \operatorname{sgn}\left(\left\langle \bar{a}_{i\ell}^{k|k}, v \right\rangle\right) = \bar{q}_{i\ell}^{k|k} \operatorname{sgn}\left(\mu_{i\ell}^{k|k} - v_n\right), \tag{23b}$$

$$\bar{p}_{i\ell}^{k|k} \left| \left\langle \bar{a}_{i\ell}^{k|k}, \hat{v} \right\rangle = \rho_{i\ell}^{k|k} \left| \mu_{i\ell}^{k|k} - v_n \right|, \right. \tag{23c}$$

where $\bar{q}_{i\ell}^{k|k} = q_{i\ell}^{k|k} \operatorname{sgn}\left(-\tilde{a}_{i\ell}^{k|k}\right)$ and $\rho_{i\ell}^{k|k} = p_{i\ell}^{k|k} \left| \tilde{a}_{i\ell}^{k|k} \right|$. Using these definitions, the inner integral of (21) is restated as

$$\int_{-\infty}^{\infty} \psi_i(v)\,dv_n = e^{j b_i^{k|k} \hat{v}} \int_{-\infty}^{\infty} g_i^{k|k} \left[\prod_{\ell=1}^{\bar{n}_{i}^{k|k}} \bar{q}_{i\ell}^{k|k} \operatorname{sgn}\left(\mu_{i\ell}^{k|k} - v_n\right) \right] \times \exp\left[\sum_{\ell=1}^{\bar{n}_{i}^{k|k}} \rho_{i\ell}^{k|k} \left| \mu_{i\ell}^{k|k} - v_n \right| + j \tilde{b}_i^{k|k} v_n \right] dv_n. \tag{24}$$

The convolution integral in (24) is shown in [13, 14] to have a closed form solution composed as a sum with $\bar{n}_{i\ell}^{k|k}$ terms, each of which is structurally similar to the terms in $\psi_i(v)$. That is, there will be a new g function which is a function of signs of inner products of $\hat{v}$.

Therefore, this integration process can be repeated until all of the integrals are taken, and a closed form solution of the conditional performance index is determined. For a vector state system of general order n, a closed form analytical solution is not attainable due to the complexity of the coefficient functions $g_i^{k|k}$. However, a closed form expression for the conditional performance index can be obtained for a two-state system, presented next, by using the structure from [17].

V. The Conditional Performance Index for a Second Order System

Now, let us limit our discussion to a second order system, i.e., $n = 2$, in order to use the structure for the cpdf's CF presented in [16]. This alternate, two state structure for the CF of the cpdf takes advantage of relationships not yet generalized to the general vector-state case that drastically reduces the number of terms needed to express the CF. In particular, there is a simpler structure for the exponential argument in (17c), as well as a simpler, closed form representation for the $g_i^{k|k}$ coefficient functions in (17a).

The structure for the CF for the ucpdf is given by

$$\bar{\phi}_{k|k}(v) = \sum_{i=1}^{N_{k|k}} G_{k|k}^{i}(v) \cdot \exp\left[-\sum_{\ell=1}^{L_{k|k}^{i}} P_i^{\ell} \left(B_{k|k}^{M_i^{\ell}} v \right)^2 + j \sum_{\ell=1}^{Z_i} B_{k|k}^{M_i^{\ell}} v \right], \tag{25}$$

which consists of a sum of $N_{k|k}$ similar terms. By using the efficient algorithm presented in [17] instead of the general vector-state algorithm from [13], the number of terms needed to express the cdpf's CF is dramatically reduced (see [17] for an explicit comparison), which also significantly improves the efficiency of the controller.

Each of these terms is a product of a coefficient function $G_{k|k}^{i}(v)$ and an exponential function whose argument involves a sum of absolute values equivalent to (17c). In the exponential argument: there is a sum of $L_i^{k|k}$ absolute value terms; the P_i^{ℓ} and Z_i^{ℓ} values represent real, scalar constants where $P_i^{\ell} > 0$; the $B_{k|k}^{M_i}$s are 1×2 row vectors, called the fundamental directions; and the M_i^{ℓ} represent integers that index the fundamental directions that multiply v. The parameters for the two-state estimator are constructed recursively in the measurement update process, as described in [16]. All of these parameters correspond with those of the structure for the cpdf's CF presented in Section IV. The P_i^{ℓ} and Z_i^{ℓ} parameters correspond to $p_{i\ell}^{k|k}$ and $b_i^{k|k}$, respectively, and the fundamental directions $B_{k|k}^{M_i}$ are the same as the $a_{i\ell}^{k|k}$ vectors.

The coefficient functions $G_{k|k}^{i}(v)$ equal the $g_i^{k|k}$ from (17a), and are rational functions of polynomials of sums of sign functions, given by

$$G_{k|k}^{i}(v) = \frac{1}{(2\pi)^n} \Bigg\{ a^i + b_i \operatorname{sgn}\left(B_{k|k}^{M_i^1} v\right) \operatorname{sgn}\left(B_{k|k}^{M_i^2} v\right) + j c_i \operatorname{sgn}\left(B_{k|k}^{M_i^1} v\right) + j d_i \operatorname{sgn}\left(B_{k|k}^{M_i^2} v\right) \times \prod_{r=1}^{L_{k|k}^{i}-2} \frac{1}{j\theta_{k-r}^{i} + \gamma + S_{k-r}^{i}(B_{k|k} v)} - \frac{1}{j\theta_{k-r}^{i} - \gamma + S_{k-r}^{i}(B_{k|k} v)} \Bigg\}, \tag{26a}$$

where

$$S_{k-r}^{i}(B_{k|k} v) = \sum_{\ell=1}^{L_{k-r|k-r}^{i}} P_i^{\ell} \left(B_{k-r|k-r}^{M_i} \right)^{\mathrm{T}} \operatorname{sgn}\left(B_{k|k}^{M_i} v \right), \tag{26b}$$

$$\theta_{k-r}^{i} = z(k-r) - Z_i^1 B_{k-r|k-r}^{M_i^1} \mathrm{H}^{\mathrm{T}} - Z_i^2 B_{k-r|k-r}^{M_i^2} \mathrm{H}^{\mathrm{T}}. \tag{26c}$$

The arguments $y^{k|k}(v)$ of $g_i^{k|k}$, given in (17b), correspond to the $S_{k-r|k-r}^{i}(B_{k|k} v)$ defined above.

The state propagation uses the same formula as in (18) and [13, 17]. This process affects the exponential argument of $\varphi_{k|k}$ by adding new terms to the real part and a transformation on the fundamental

directions $B^{M_i}_{k|k}$. The coefficient functions remain unchanged as polynomials of sign functions, but the arguments of the sign functions have the same transformation as in the exponential argument, so that $G^i_{p|k}(v) = G^i_{k|k}(\Phi^{mT} v)$. The state propagation does not add any new terms to the sum, so that $N_{p|k} = N_{k|k}$. Hence, the m-step state propagated cpdf's CF is given by

$$\bar{\phi}_{p|k}(v) = \prod_{i=1}^{N_{p|k}} G^i_{p|k}(v) \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P_i B^{M_i}_{p|k} v + j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{p|k}\right) v\right]. \tag{27}$$

Using this expression for the unnormalized CF, we can derive an expression for the performance index in closed form. The transformed objective function in the performance index (14) is given by

$$L^*_{x(p)}(v) = e^{-\eta_{p,1}|v_1| + j\bar{x}_1 v_1 - \eta_{p,2}|v_2| + j\bar{x}_2 v_2} \tag{28}$$

where, since k will be a constant through this process, the time subscript of $\bar{x}(p)$ is replaced with element subscripts as $\bar{x}(p) = [\bar{x}_1 \quad \bar{x}_2]^T$. The 2-state specific version of (15) then becomes

$$\begin{aligned} J_{Z_k} &= \frac{M}{(2\pi)^2} \int_{-\infty}^{\infty} L^*_{x(p)}(v_1, v_2) \cdot \varphi_{Z_k}(v_1, v_2)\, dv_1 v_2 \\ &= \frac{M}{(2\pi)^2 f_{Z_k}} \int_{-\infty}^{\infty} \exp\left(-\eta_1|v_1| - \eta_2|v_2| + j\bar{x}_1 v_1 + j\bar{x}_2 v_2\right) \\ &\quad \times \prod_{i=1}^{N_{p|k}} G^i_{p|k}(v) \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P_i B^{M_i}_{p|k} v + j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{p|k}\right) v\right] dv. \end{aligned} \tag{29}$$

The integral with respect to v_2 is now taken. We use a second subscript to denote the individual elements of the M_i row of $B_{p|k}$ as $B^{M_i}_{p|k} = B^{M_i}_{k,1} \quad B^{M_i}_{k,2}$. This allows us to decompose the complex part of the exponential argument as

$$j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{p|k}\right) v = j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{k,1}\right) v_1 + j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{k,2}\right) v_2, \tag{30}$$

and rewrite J_{Z_k} as

$$\begin{aligned} J_{Z_k} &= \frac{1}{(2\pi)^2 f_{Z_k}} \prod_{i=k}^{p-1} \frac{\zeta/\pi}{u^2(i) + \zeta_i^2} \\ &\quad \times \prod_{i=1}^{N_{p|k}} \int_{-\infty}^{\infty} \exp\left(-\eta_1|v_1| + j\bar{x}_1 v_1 + j \sum_{\ell=1}^{2} \left(Z_\ell B^{M_i}_{k,1}\right) v_1\right) \\ &\quad \times \int_{-\infty}^{\infty} G^i_{p|k}(v) \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P_i B^{M_i}_{p|k} v - \eta_2|v_2| \right. \\ &\quad \left. + j \left(\sum_{\ell=1}^{2} Z_\ell B^{M_i}_{k,2}\right) v + j\bar{x}_2 v_2\right] dv_2\, dv_1. \end{aligned} \tag{31}$$

Denote the i^{th} inner integral with respect to v_2 in (31) as $I_{i,2}$:

$$I_{i,2} = \int_{-\infty}^{\infty} G^i_{p|k}(v_1, v_2) \times \exp\left[-\sum_{\ell=1}^{L^i_{p|k}} P_i B^{M_i}_{k,1} v_1 + B^{M_i}_{k,2} v_2 - \eta_2|0 + v_2| + j\left(\sum_{\ell=1}^{2} Z_\ell B^{M_i}_{k,2}\right) v_2 + j\bar{x}_2 v_2\right] dv_2. \tag{32}$$

The integral $I_{i,2}$ is over the v_2 variable, but it also contains v_1 in the absolute value terms in the argument of the exponential and in the coefficient function $G^i_{p|k}$. In order to solve it, we need to use the integral of absolute values method presented in [13, 14, 16]. This method involves defining a set of scalar constants ρ_ℓ and ξ^i_k, as well as scalar variables μ_ℓ that depend on v_1, and thus are constants in this integration.

If $B^{M_i}_{k,2} = 0$, then when this constant multiplies v_2 the variable disappears, and that term comes out of the integral. This will affect the specific form of second integral over v_1, but the method presented here can still be used. However, for simplicity of presentation, in this derivation we assume that all the $B^{M_i}_{k,2} \neq 0$. Then, we can construct a $\{\mu_\ell\}_1^{L^i_{p|k}+1}$ set that transforms $I_{i,2}$ into the integral of absolute values structure. This assumption does not affect generality, because if one of the $B^{M_i}_{k,2}$ equals zero for some i, then that absolute value term would not be a function of v_2 and would come out of $I_{i,2}$ and integrated later.

The set of variables, denoted $\{\mu_\ell\}$, is constructed as

$$\begin{aligned} B^{M_i}_{k,1} v_1 + B^{M_i}_{k,2} v_2 &= -B^{M_i}_{k,2} \left[\frac{-B^{M_i}_{k,1} v_1}{B^{M_i}_{k,2}} - v_2\right] \\ &= -B^{M_i}_{k,2} (\mu_\ell - v_2) \end{aligned} \tag{33a}$$

where

$$\mu_\ell = \begin{cases} \dfrac{-B^{M_i}_{k,1} v_1}{B^{M_i}_{k,2}} & R \in \{1, \ldots, L^i_{p|k}\} \\ 0 & R = L^i_{p|k} + 1. \end{cases} \tag{33b}$$

Similarly, for the argument of the exponential, we can construct a set of $\{\rho_\ell\}_1^{L^i_{p|k}+1}$ as

$$\rho_\ell = \begin{cases} -B^{M_i}_{k,2} & R \in \{1, \ldots, L^i_{p|k}\} \\ \eta_2 & R = L^i_{p|k} + 1 \end{cases} \tag{33c}$$

and a scalar number ξ^i_k as

$$\xi^i_k = \bar{x}_2 + \sum_{\ell=1}^{2} Z_\ell B^{M_i}_{k,2}. \tag{33d}$$

The solution to an integral of an exponent of absolute values requires dividing the domain of integration into regions in which the integrand is continuous. Since $G^i_{p|k}(v)$ is piecewise-constant, its discontinuities lie on the boundaries of these regions, and hence $G^i_{p|k}(v)$ is treated as a constant in each integral. In order for the $G_{p|k}(v)$ coefficients to be consistent with the form in (37), a procedure mirroring that of the measurement update process from [16, 17] must be used. This involves rewriting the first bracket term in (26a) as a sum of sign functions, using the identity $\text{sgn}\left(B^{M_i}_{k|k} v\right) \text{sgn}\left(B^{M_i}_{k|k} v\right) = \frac{1}{2}\left[\text{sgn}\left(B^{M_i}_{k|k} v\right) + \text{sgn}\left(B^{M_i}_{k|k} v\right)\right]^2 - 1$ and the substitutions

$$\bar{\rho}_1 = c_i \text{sgn}\left(-B^{M_i}_{k,2}\right), \quad \bar{\rho}_2 = d_i \text{sgn}\left(B^{M_i}_{k,2}\right), \quad \bar{\rho}_\ell = 0 \text{ for } R > 2,$$

$$\tilde{\rho}_1 = 1, \qquad \tilde{\rho}_2 = 1, \qquad \tilde{\rho}_{\mathcal{E}} = 0 \text{ for } R > 2. \tag{34}$$

Furthermore, for the rest of the bracket terms in (26a) we can rewrite the sums given by S_{k-r} in (26b) by defining constants ${}^{(r)}\check{\rho}_\ell$ as

$$ {}^{(r)}\check{\rho}_\ell = \begin{cases} P^{\ell}\left(B_{k-r|k-r}^{M_i} H^{\mathrm{T}}\right) \operatorname{sgn}\left(-B_{k,2}^{M_i}\right), & R = 1, \ldots, L_{p|k}^i - r \\ 0 & R = L_{p|k}^i - r + 1, \ldots, L_{p|k}^i + 1 \end{cases} \tag{35} $$

in order to write $G_{p|k}^i$ in (32) as

$$ \begin{aligned} G_{p|k}^i(v - H^{\mathrm{T}}\sigma) &= \\ &G_{p|k}^i\Big(\sum_{\ell=1}^{L_{p|k}^i+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \sum_{\ell=1}^{L_{p|k}^i+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \ldots \\ &\ldots, \sum_{\ell=1}^{L_{p|k}^i+1} {}^{(L_{p|k}^i-2)}\check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma), \ldots, \sum_{\ell=1}^{L_{p|k}^i+1} {}^{(1)}\check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\Big) \\ &:= G_{p|k}^i\Big(\sum_{\ell=1}^{L_{p|k}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\Big). \end{aligned} \tag{36} $$

Here, the variable $\check{\rho}_\ell$ is a shorthand variable representing all of the ${}^{(r)}\check{\rho}_\ell$, $\tilde{\rho}_\ell$, and $\bar{\rho}_\ell$ constants in all of the sums in $G_{p|k}^i$.

Then, to write this integral in a form consistent with the measurement update formula, let $\sigma := v_2$ in order to write the one-dimensional integral as

$$ I_{i,2} = \int_{-\infty}^{\infty} G_{p|k}^i\Big(\sum_{\ell=1}^{L_{p|k}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \sigma)\Big) \times \exp\Big(-\sum_{\ell=1}^{L_{p|k}^i+1} \rho_\ell |\mu_\ell - \sigma| + j\xi_k^i \sigma\Big) d\sigma. \tag{37} $$

The solution is given as a sum of $L_{p|k}^i + 1$ terms as

$$ \begin{aligned} I_{i,2}(v_1) = &\sum_{m=1}^{L_{p|k}^i+1} \exp\Big(-\sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \rho_\ell |\mu_\ell - \mu_m| + j\xi_k^i \mu_m\Big) \\ &\times \frac{G_{p|k}^i\Big(+\check{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\Big)}{j\xi_k^i + \rho_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)} \\ &- \frac{G_{p|k}^i\Big(-\check{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\Big)}{j\xi_k^i - \rho_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \rho_\ell \operatorname{sgn}(\mu_\ell - \mu_m)}. \end{aligned} \tag{38a} $$

where

$$ \begin{aligned} G_{p|k}^i\Big(\pm\check{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \check{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\Big) = &\\ \Big(a_i + b_i\Big(\frac{1}{2} \pm \tilde{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \tilde{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\Big)^2 - 1\Big) & \\ + j\Big(\pm\bar{\rho}_m + \sum_{\substack{\ell=1 \\ \ell \neq m}}^{L_{p|k}^i+1} \bar{\rho}_\ell \operatorname{sgn}(\mu_\ell - \mu_m)\Big) & \\ \times \prod_{p=1}^{L_{k|k}^i-2} \frac{1}{2\pi} \frac{1}{j\theta_{k-r}^i + \gamma + S_{k-r}^i(\pm{}^{(r)}\hat{\rho}_m, \mu_\ell - \mu_m)} & \\ - \frac{1}{j\theta_{k-r}^i + \gamma + S_{k-r}^i(\pm{}^{(r)}\hat{\rho}_m, \mu_\ell - \mu_m)}, & \end{aligned} \tag{38b} $$

and for $r \in \{1, \ldots, L_{k|k}^i - 2\}$,

$$ \begin{aligned} &S_{k-r}^i(\pm{}^{(r)}\hat{\rho}_m, \mu_\ell - \mu_m) \\ &= \pm P^{M_i^m} B_{k-r|k-r}^{M_i^m} H^{\mathrm{T}} \\ &\quad \sum_{\ell=1}^{L_{k-r|k-r}^i} P^{\ell}\left(B_{k-r|k-r}^{M_i} H^{\mathrm{T}}\right) \operatorname{sgn}\left(-B_{k,2}^{M_i}\right) \operatorname{sgn}(\mu_\ell - \mu_m) \\ &= \pm P_i^{m} B_{k-r|k-r}^{M_i^m} H^{\mathrm{T}} \\ &\quad \sum_{\ell=1}^{L_{k-r|k-r}^i} P_i^{\ell}\left(B_{k-r|k-r}^{M_i} H^{\mathrm{T}}\right) \operatorname{sgn}\left(-B_{k,2}^{M_i}\right) - \frac{B_{k,1}^{M_i^m}}{B_{k,2}^{M_i^m}} + \frac{B_{k,1}^{M_i^m}}{B_{k,2}^{M_i^m}} \\ &\quad \times \operatorname{sgn}(v_1). \end{aligned} \tag{38c} $$

In the two-state estimator [17], algebraic relationships are used to reduce the complicated bracket terms produced by the update process into simpler, polynomial forms. Similarly, those same algebraic relationships are used here to simplify (38a). Denote the coefficient of the m^{th} term in the sum as $G_{p|k}^{i,m}$ and consider the numerators of the coefficient function in the brackets in (38a) given by (38b). This complicated term can be reduced to a simple function of $\operatorname{sgn}(v_1)$ by rewriting it in the form of the left hand side of (48a), and then applying the algebraic identity presented in the Appendix. This process begins by rewriting the numerator functions (38b) in a four parameter form involving only $\operatorname{sgn}(v_1)$ as

$$ a_i^m \pm b_i^m \cdot \operatorname{sgn}(v_1) \pm j c_i^m + j d_i^m \cdot \operatorname{sgn}(v_1). \tag{39} $$

Then, the real part of the exponential argument in (38a) and the sums in the denominators of the coefficient function in (38b) can be reduced to a single constant times $|v_1|$ or $\operatorname{sgn}(v_1)$, respectively,

where those constants are defined as

$$D_{i,m}|v_1| = \sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \rho_\ell |\mu_\ell - \mu_m| = \left(\sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \rho_\ell \left(\frac{-B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} - \frac{-B_{k,1}^{M_i^m}}{B_{k,2}^{M_i^m}} \right) \right) |v_1|, \quad (40a)$$

$$\check{D}_{i,m}\mathrm{sgn}(v_1) = \sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \rho_\ell \mathrm{sgn}(\mu_\ell - \mu_m) = \left(\sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \rho_\ell \mathrm{sgn}\left(\frac{-B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} - \frac{-B_{k,1}^{M_i^m}}{B_{k,2}^{M_i^m}} \right) \right) \mathrm{sgn}(v_1). \quad (40b)$$

Expressing the numerators of (38b) as (39), and using $\check{D}_{i,m}$, then the result in the Appendix can be used to express the coefficient function as a two parameter form given in (48a), where those parameters are denoted $a_{i,m}$ and $d_{i,m}$. For the exponential argument, use $D_{i,m}$ and rewrite the complex part as

$$\hat{\xi}_k^i v_1 = \left(\xi_k^i \cdot \frac{-B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} \right) \cdot v_1 \quad (41)$$

in order to express (38a) in a much simpler form given by

$$I_{i,2}(v_1) = \prod_{m=1}^{L_{p/k+1}^{i}} \{a_{i,m} + j\, d_{i,m}\mathrm{sgn}(v_1)\} \cdot e^{-D_{i,m}|v_1| + j\,\hat{\xi}_k^i v_1}. \quad (42)$$

To obtain the form in (39) for (38b), a procedure mirroring that of the estimator update process is applied, which uses the algebraic relationships presented in the Appendix of [17]. For $m = 1$ and $m = 2$, the numerators of (38a) are not equal and hence cannot come out of the bracket term. Then, the numerators for $m = 1$ are given by

$$\begin{aligned} &a_i + b_i \left(\frac{1}{2} \left(\pm \tilde{\rho}_m + \sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \tilde{\rho}_\ell \mathrm{sgn}(\mu_2 - \mu_1) \right)^2 - 1 \right) \\ &\quad + j \left[\pm \bar{\rho}_1 + \sum_{\ell=2}^{L_{p/k+1}^{i}} \bar{\rho}_\ell \mathrm{sgn}(\mu_2 - \mu_1) \right] \\ &= a_i \pm b_i \mathrm{sgn}\left(B_{k,2}^{M_i^1} \cdot B_{k,2}^{M_i^2} \right) \mathrm{sgn}(\mu_2 - \mu_1) \\ &\quad \pm j c_i \mathrm{sgn}\left(B_{k,2}^{M_i^1} \right) + j d_i \mathrm{sgn}\left(B_{k,2}^{M_i^2} \right) \mathrm{sgn}(\mu_2 - \mu_1) \\ &= a_i \pm b_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} + \frac{B_{k,1}^{M_i^2}}{B_{k,2}^{M_i^2}} \right) \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) \\ &\quad + j c_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) + d_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^2}}{B_{k,2}^{M_i^2}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) \\ &\quad \times \mathrm{sgn}(v_1). \end{aligned} \quad (43)$$

The same manipulations can be done for the numerators for $m = 2$ by interchanging μ_1 and c_i with μ_2 and d_i, respectively. The + and − in (43) refer to the left and right numerators, respectively, in (38a) (38b) by substituting in (38c). This product is given in Result 2 from the Appendix of [17] and produces a term of the same form as (43). Hence, this result can be combined with the next bracket term in the product in (38c), eventually expressing both numerators as (39).

For $m > 2$, this first bracket term produces a simpler form due to two sign functions canceling each other out. Recalling that $\bar{\rho}_\ell = \rho_\ell = 0$ for $R > 2$, the new term obtained for $m = 3$ corresponds to the coefficient

$$\begin{aligned} &a_i + b_i \left(\frac{1}{2} \left(\pm \tilde{\rho}_m + \sum_{\substack{\ell=1 \\ \ell\neq m}}^{L_{p/k+1}^{i}} \tilde{\rho}_\ell \mathrm{sgn}(\mu_1 - \mu_3) \right)^2 - 1 \right) \\ &\quad + j \left[\pm \bar{\rho}_1 + \sum_{\ell=2}^{L_{p/k+1}^{i}} \bar{\rho}_\ell \mathrm{sgn}(\mu_2 - \mu_3) \right] \\ &= a_i + b_i \mathrm{sgn}\left(B_{k,2}^{M_i^1} \cdot B_{k,2}^{M_i^2} \right) \mathrm{sgn}(\mu_1 - \mu_3) \mathrm{sgn}(\mu_2 - \mu_3) \\ &\quad + j c_i \mathrm{sgn}\left(B_{k,2}^{M_i^1} \right) \mathrm{sgn}(\mu_1 - \mu_3) + j d_i \mathrm{sgn}\left(B_{k,2}^{M_i^2} \right) \mathrm{sgn}(\mu_2 - \mu_3) \\ &= a_i + b_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^2}}{B_{k,2}^{M_i^2}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) \\ &\quad + j \left[c_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^1}}{B_{k,2}^{M_i^1}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) + d_i \mathrm{sgn}\left(-\frac{B_{k,1}^{M_i^2}}{B_{k,2}^{M_i^2}} + \frac{B_{k,1}^{M_i^3}}{B_{k,2}^{M_i^3}} \right) \right] \\ &\quad \times \mathrm{sgn}(v_1). \end{aligned} \quad (44)$$

Now, substitute (42) into (31) and denote the outer integral with respect to v_1 by

$$\begin{aligned} I_{i,1} &= \prod_{m=1}^{L_{p/k+1}^{i}} \{a_{i,m} + j\, d_{i,m}\mathrm{sgn}(v_1)\} \\ &\quad \times \int_{-\infty}^{\infty} \exp\left(-D_{i,m}|v_1| + j\hat{\xi}_k^i v_1 \right) \exp(-\eta_1 |v_1| + j\bar{x}_1 v_1) \\ &\quad \times \exp\left(j \sum_{\ell=1}^{2} Z_i^\ell B_{k,1}^{M_i^1} v_1 \right) dv_1 \\ &= \prod_{m=1}^{L_{p/k+1}^{i}} \{a_{i,m} + j\, d_{i,m}\mathrm{sgn}(v_1)\} \\ &\quad \times \int_{-\infty}^{\infty} \exp\left(-\left(D_{i,m} + \eta_1 \right) |v_1| + j \left(\hat{\xi}_k^i + \bar{x}_1 + \sum_{\ell=1}^{2} Z_i^\ell B_{k,1}^{M_i^1} \right) v_1 \right) dv_1. \end{aligned} \quad (45)$$

This integral has a form identical to the measurement update process for a scalar system [10]. Its solution is given by

$$I_{i,1} = \prod_{m=1}^{L_{p/k+1}^{i}} \frac{a_{i,m} - j\, d_{i,m}}{\sqrt{2\pi}} \cdot \left[\frac{1}{D_{i,m} + \eta_1 - j\left(\hat{\xi}_k^i + \bar{x}_1 + \sum_{\ell=1}^{2} Z_i^\ell B_{k,1}^{M_i^1} \right)} + \frac{1}{D_{i,m} + \eta_1 + j\left(\hat{\xi}_k^i + \bar{x}_1 + \sum_{\ell=1}^{2} Z_i^\ell B_{k,1}^{M_i^1} \right)} \right].$$

$$\sum_{m=1}^{2}\Bigg(\hat{\xi}_k+\bar{x}_1+\sum_{\ell=1}^{}Z_i^{\ell}B_{k,1}^{M'}+D_{i,m}+\eta_1\Bigg)^2. \tag{46}$$

that are given by (38b).

Hence, since the first bracket term in (38b) can be expressed as (43), it can be combined with a bracket term from the product in

Finally, using (46) in (31) and reintroducing the time dependence notation, the conditional performance index is given by

$$J_{Z_k}(p) = \frac{1}{2\pi^2 f_{Z_k}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=k}^{k+1} \frac{\zeta_i/\pi}{u^2(i) + \zeta_i^2} \prod_{i=1}^{2} \prod_{m=1}^{2} \left| a_{i,m} D_{i,m} + \eta_{p,1} - d_{i,m} \left[\bar{x}_1(p) + \sum_{\ell=1}^{M_i} Z_i B_{k,1}^{\ell} \right] \right|^{-2} \cdot \left| \hat{c}_k + \bar{x}_1(p) + \sum_{\ell=1}^{M_i} Z_i B_{k,1}^{\ell} + \frac{D_{i,m}}{2} + \eta_{p,1} \right|^{-2} \quad (47)$$

This closed form conditional performance index was observed to be non-convex and depends on the control input sequence $\{u(k), \ldots, u(p-1)\}$ in a complex way. Specifically, $\bar{x}_1(p) \;\; \bar{x}_2(p)^{\mathrm{T}} = \bar{x}(p)$ depends on the control sequence, along with the parameters $a_{i,m}$ and $d_{i,m}$ that also depend on the measurement sequence $\tilde{Z}_k$. Therefore, we maximize (47) numerically. The optimization is done in two steps: first, the global optimum of the conditional performance index, expressed as a double sum in (47), is optimized with respect to the control input with a coarse gradient search in order to find an approximate optimal control sequence. Then, we use the accelerated gradient search method [21] starting from that approximate sequence in order to refine the optimal control solution.

VI. Numerical Examples

Here, we present two sets of examples, the first of which shows the optimal control versus the measurement for the first time step only, and the second set shows two multi-step examples. All of these examples use a two-step horizon, i.e. $m = 2$, so that there exists a control sequence that can drive our two-state system to the origin over this horizon length. However, as we are using model predictive control, only the first control input of this sequence is applied at that time step. Gaussian parameters used for the LEG and for Gaussian noises are closest in the $S\alpha S$ sense to their corresponding Cauchy distributions. For clarity of presentation in the figures, the first update occurs at $k = 0$ instead of $k = 1$. Furthermore, since the control weightings remain constant across time steps, we will drop the time notation from the parameters η and ζ.

All the simulations use the same system dynamics with $\mathsf{H} = [1 \;\; 1]$, $\Gamma^{\mathrm{T}} = [0.5 \;\; 1]$, $\Lambda^{\mathrm{T}} = [0.5 \;\; 1]$, $m = 2$, and the eigenvalues of Φ are $0.8 \pm 0.55j$. The terminal state weightings are $\eta = [1 \;\; 1]$, and when control weightings are used they are $\zeta = 10$. The initial condition's scaling parameters are given by $\alpha = [0.8 \;\; 0.8]$, and the process and measurement noise parameters β and γ are either 0.5 or 0.1.

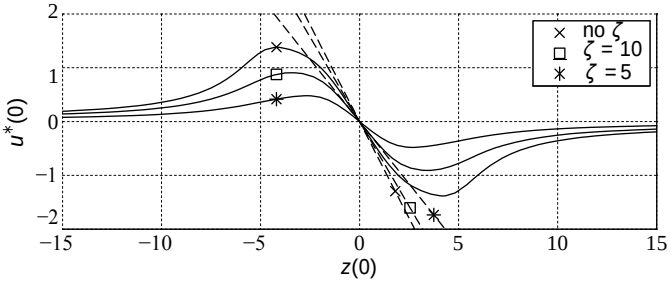
All of our examples compare our Cauchy optimal model predictive controller with a similar LEG model predictive optimal controller [1]. The LEG estimator assumes that the stochastic inputs are described by the Gaussian pdfs that are closest, in a least squared sense, to the given Cauchy pdfs. The LEG controller assumes that its objective functions of the state and control resemble scaled Gaussian pdfs that are closest, in the least-squared sense, to the scaled Cauchy pdfs in (10). The LEG design details can be found in [12–14, 16, 17], where similar comparisons and least square fits were used. The LEG controllers' responses are shown in dashed lines in the figures.

A. One-Step Control Examples

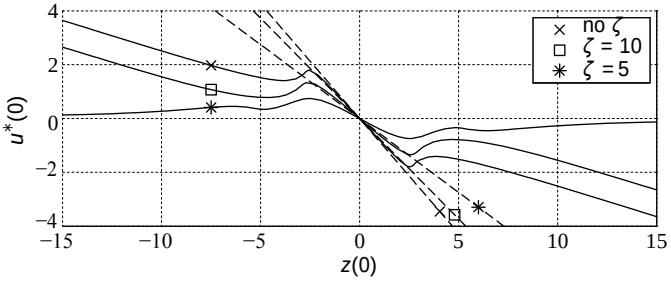
The first set of examples are shown in Fig. 1. These figures show the applied optimal control input at the first time step given the first measurement. In the two cases presented, all the systems parameters are the same, except in Fig. 1(a) $\gamma > \beta$ (i.e., more measurement than state uncertainty), and in Fig. 1(b) $\beta > \gamma$ (i.e., more state than measurement uncertainty).

The example in Fig. 1(a) shows that the Cauchy controller is nearly linear for small measurements and reduces its control effort to zero as the measurement deviations become large. This is in contrast to the LEG controller, which is linear and thus responds strongly to large measurement deviations. This behavior in the Cauchy controller occurs when the measurement uncertainty is larger than the state uncertainty. In the opposite case shown in Fig. 1(b), the measurement has less uncertainty than the state. Here, the Cauchy controller's response closely matches that of the LEG in a neighborhood of the origin, and in fact responds even more strongly than the LEG for large measurement deviations.

The three different curves in both of these figures depict the control signals for three different control weights: no control weight, $\zeta = 10$, and $\zeta = 5$. As expected, heavier control weights (i.e. smaller ζ) reduce the control effort. Even without any control weighting, the response in Fig. 1(a) goes to zero for large measurement deviations. The fact that this behavior is seen when there is no control weighting implies that the attenuation of the control signal for large measurement deviations is due to the cpdf and not the objective function. Moreover, this behavior is not shared by the LEG controller that uses a similar objective function but assumes light-tailed, Gaussian distributions.



(a) $\beta = 0.1$, $\gamma = 0.5$.



(b) $\beta = 0.5$, $\gamma = 0.1$.

Fig. 1. Optimal control vs the measurement for the first time step.

B. Multi-Step Control Examples

The complexity of evaluating the conditional performance index grows as the number of terms increases across time steps, as indicated in (47). For implementable control, this growth needs to be arrested. The full information CF of the ucpdf (25) is approximated by a CF of a ucpdf conditioned on a fixed sliding window of the most recent measurements, as described in [17], where the number of measurements in this horizon is denoted N_Z. Hence, the first N_Z measurement updates are performed normally, and the control optimization is performed using this CF. Then, for time steps $k > N_Z$, we initialize a new finite horizon (FH) estimator and perform N_Z measurement updates over the fixed window $\{z(k - N_Z + 1), \ldots, z(k)\}$. The conditional performance index (15) is evaluated using this FH CF.

The remaining examples in Figs 2, 3, 5, 4, and 6 are all multistep examples over 100 measurements, and all use a horizon length of $N_Z = 8$. The variations between the simulations are in the stochastic parameters, alternating which of process or measurement noise dominates the other and whether the noises are Gaussian or Cauchy.

In Fig. 2, there is more uncertainty in the state process noise than in the measurement noise. When large measurement deviations occur (such as at $k = 52$), the Cauchy controller's effort is very small even though there is no weighting on the control inputs. In contrast, the LEG controller responds with a large control effort that drives the states from their regulated state of zero. However, when large process noise inputs occur, the state deviates and the Cauchy controller applies a larger control effort than the LEG, thus regulating the state more effectively. It is interesting to compare Fig. 2 with Fig. 3 in light of Fig. 1. They suggest that when the measurement noise density parameter dominates the process noise density parameter in constructing the Cauchy controller, the effect of measurement outliers is mitigated, while still responding to state deviations due to process noise.

On the other hand, in Fig. 3 there is more uncertainty in the state than in the measurement, and the Cauchy controller behaves very much like the linear LEG controller. The state trajectories and control inputs of the Cauchy and LEG controllers appear equal, but actually their differences are much smaller than the scale of the axis and cannot be seen. Similar behavior was observed in Fig. 1(b), suggesting that when the stochastic parameters allow it, the Cauchy controller follows the measurement more. It responds in a more linear fashion to the measurements, similar to the performance of the LEG controller in that setting.

This behavior is seen again when both controllers face Gaussian noises, as in Figure 4. Here, the Cauchy controller closely follows both the control and state trajectories of the LEG, which is the true optimal solution. This demonstrates that the Cauchy controller is robust under non-impulsive noise environments, as it closely approximates the true optimal solution given by the LEG.

Fig. 5 shows two examples using control weighting $\zeta = 10$, where otherwise Fig. 5(a) has the same parameters as Fig. 2, and similarly with Figs. 5(b) and 3. Applying control weighting appears to slow down the reaction from the Cauchy controller, as seen in the response to the initial process noise pulses in Fig. 2. In Fig. 5(a) with control weighting, a one step delay is observed in the Cauchy controller's response before applying the regulating control input. This behavior appears again in Fig. 5(b), where the Cauchy controller is very delayed in applying control, leading to much larger deviations of the state than the LEG allowed. This suggests that excessive control weighting can adversely affect the regulation performance of the Cauchy controller in response to process noise pulses. However, in a situation where the measurement noise dominates, the control weighting haS a much smaller effect, and improves performance in ignoring measurement outliers.

The final set of results in Fig. 6 show situations where either measurement or process noise is Cauchy and the other is Gaussian. In Fig. 6(a) the measurement noise is Cauchy and the process noise is Gaussian. Hence, there are no large state deviations, and the Cauchy controller performs similarly to the example in Fig. 2, effectively ignoring the measurement deviations. In Fig. 6(b), the measurement noise is Gaussian and the process noise is Cauchy, and the Cauchy controller performs similarly to the example in Fig. 3. Here, the large deviations occur in the state, but there are no large measurement deviations. The Cauchy controller's performance closely matches that of the LEG in this case, except for the early time steps.

VII. Conclusions

An optimal stochastic controller was derived for vector-state, linear, discrete-time systems with additive process and measurement Cauchy distributed noises. Since the Cauchy distribution has an undefined mean and an infinite second moment, we cannot use standard objective functions, e.g., the expected value of a quadratic function of state and control variables. Therefore, a new and computable objective function was defined. As opposed to previous work, the CF of the cpdf of the state given the measurement history and the Parseval's identity are used to express the conditional performance index in a closed form. This closed-form conditional performance index is optimized numerically using an accelerated gradient search. Examples are presented that show how the proposed vector state Cauchy controller compares against an equivalent LEG controller, demonstrating the Cauchy controller's performance and improved robustness over its Gaussian counterpart.

Appendix

The term given by

$$\left(\frac{a_i^m + b_i^m \cdot \operatorname{sgn}(v_1) + jc_i^m + jd_i^m \cdot \operatorname{sgn}(v_1)}{j\theta + \rho_m + \check{D} \cdot \operatorname{sgn}(v_1)} \cdot \frac{a_i^m - b_i^m \cdot \operatorname{sgn}(v_1) - jc_i^m + jd_i^m \cdot \operatorname{sgn}(v_1)}{j\theta - \rho_m + \check{D} \cdot \operatorname{sgn}(v_1)}\right)^{-1}$$

$$= a_{i,m} + jd_{i,m} \cdot \operatorname{sgn}(v_1) \quad (48a)$$

where the parameters $a_{i,m}$ and $d_{i,m}$ are generated from the other parameters by the linear relationships

$$\begin{bmatrix} a_{i,m} \\ d_{i,m} \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \bar{a} & \bar{b} & 0 & 0 \\ -\bar{b} & \bar{a} & 0 & 0 \\ 0 & 0 & \bar{a} & \bar{b} \\ 0 & 0 & -\bar{b} & \bar{a} \end{bmatrix} \cdot \begin{bmatrix} \rho_m & -\check{D} & -\theta & 0 \\ \check{D} & \rho_m & 0 & -\theta \\ \theta & 0 & \rho_m & -\check{D} \\ 0 & \theta & \check{D} & \rho_m \end{bmatrix} \cdot \begin{bmatrix} a_i^m \\ b_i^m \\ c_i^m \\ d_i^m \end{bmatrix}, \quad (48b)$$

where

$$\bar{a} = \check{D}^2 - \rho_m^2 - \theta^2, \qquad \bar{b} = -2\theta\rho_m,$$
$$\Delta = \pi \cdot \left(\bar{a}^2 + \bar{b}^2\right). \quad (48c)$$

References

[1] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

[2] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

[3] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. W. Rayner, "Near optimal detection of signals in impulsive noise modeled with asymmetric alpha-stable distribution," *IEEE Communications Letters*, vol. 2, no. 10, pp. 282–284, Oct. 1998.

[4] P. Reeves, "A non-gaussian turbulence simulation," Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-69-67, 1969.

[5] P. Zarchan, *Tactical and Strategic Missile Guidance*, 6th ed. Reston, VA: American Institute of Aeronautics and Astronautics, Inc., 2012.

[6] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gassian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall, 1994.

[7] P. Tsakalides and C. L. Nikias, *Deviation from Normality in Statistical Signal Processing: Parameter Estimation with Alpha-Stable Distributions;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

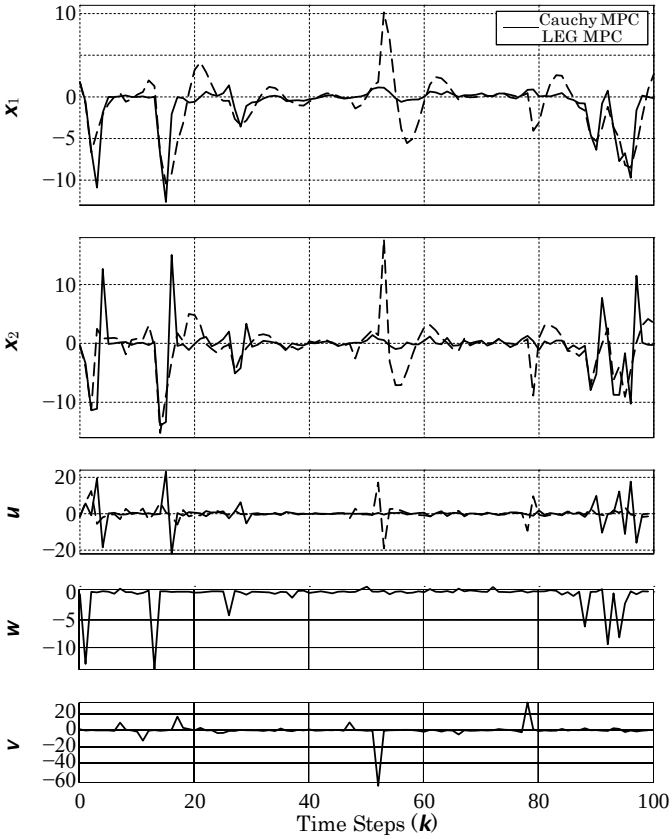
[8] G. A. Tsihrintzis, *Statistical Modeling and Receiver Design for Multi-User Communication Networks;* in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser, 1998.

[9] G. A. Hewer, R. D. Martin, and J. Zeh, "Robust preprocessing for kalman filtering of glint noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 120–128, January 1987.

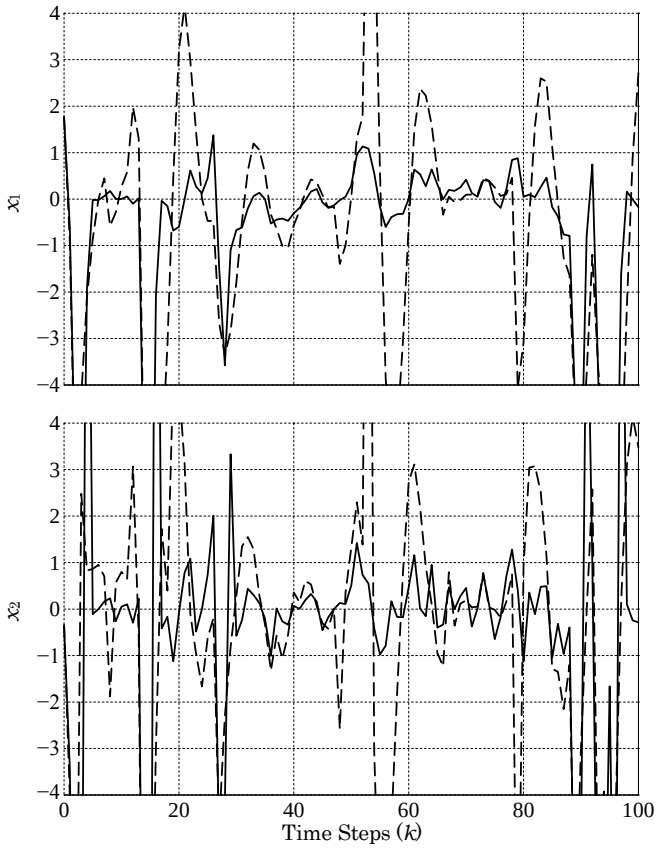
[10] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1329–1342, 2010.

Control for a Scalar Linear System with Additive Cauchy Noise," in

[12] ——, "A Stochastic Controller for a Scalar Linear System with Additive Cauchy Noise," *Automatica*.
[13] M. Idan and J. L. Speyer, "Multivariate cauchy estimator with scalar measurement and process noises," *Submitted to SIAM*, 2013.
[14] ——, "Multivariate Cauchy Estimator with Scalar Measurement and Process Noises," in *Proceedings of the* 52nd *IEEE Conference on Decision and Control*, Florence, Italy, December 2013.
[15] ——, "State estimation for linear scalar dynamic systems with additive cauchy noises: Characteristic function approach," *SIAM J. Control Optim.*, vol. 50, no. 4, pp. 1971–1994, 2012.
[16] J. L. Speyer, M. Idan, and J. Fernández, "The Two-State Estimator for Linear System with Additive Measurement and Process Cauchy Noise," in *IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2012.
[17] J. H. Fernandez, J. L. Speyer, and M. Idan, "Linear Dynamic Systems with Additive Cauchy Noises - Part 1: Stochastic Estimation for Two-State Systems," *Submitted to IEEE Transactions*, 2013.
[18] J. Fernández, J. L. Speyer, and M. Idan, "A Stochastic Controller for Vector Linear Systems with Additive Cauchy Noises," in *Proceedings of the* 52nd *IEEE Conference on Decision and Control*, Florence, Italy, December 2013.
[19] M. Morari and E. Zafiriou, *Robust Process Control*. New Jersey: Prentice-Hall, 1989.
[20] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. New York: Springer-Verlag, 1975, ch. 5 & 6.
[21] R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," *The Computer Journal*, vol. 6, no. 2, pp. 163–168, 1963.

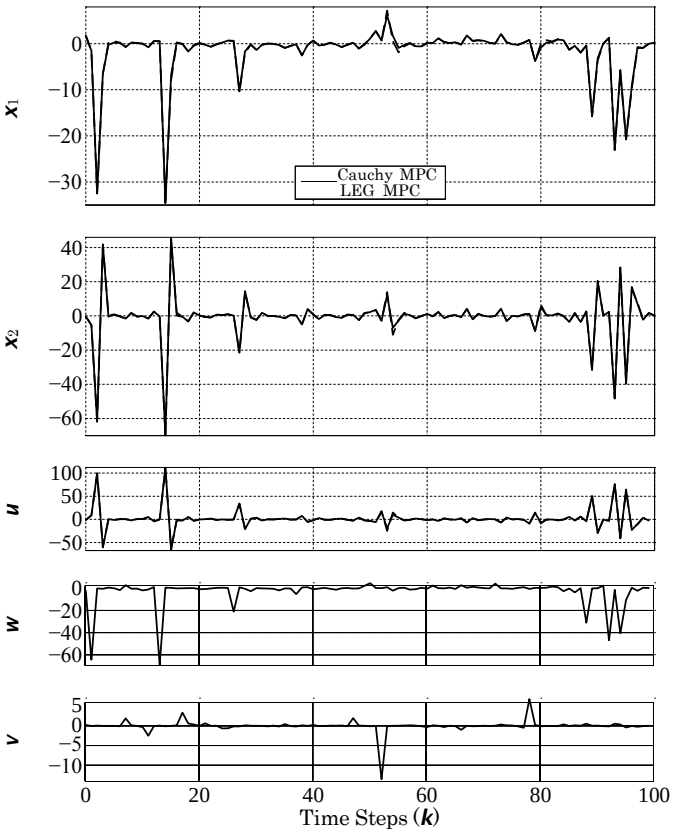


(a) Full view.

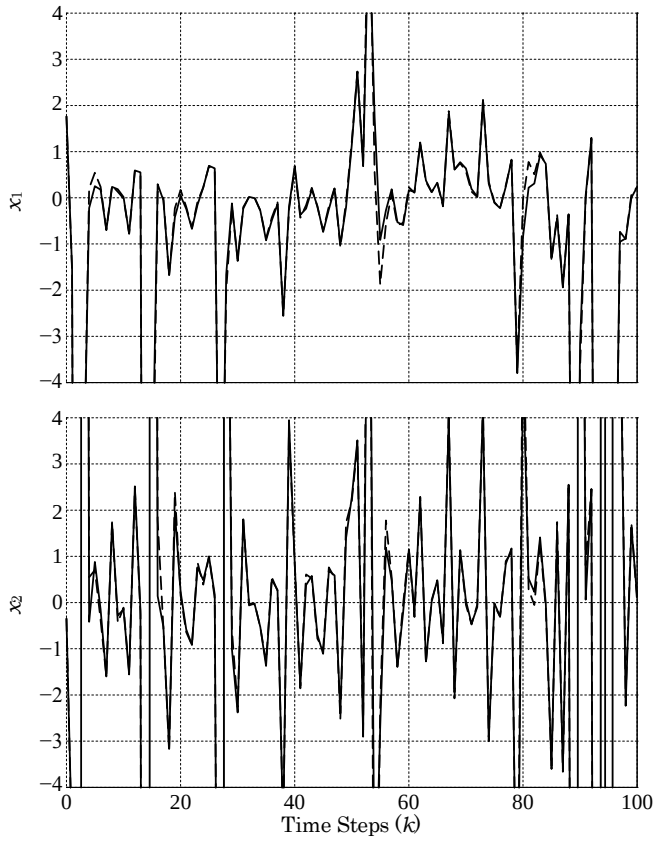


(b) Zoomed view.

Fig. 2. Cauchy and LEG controllers' performance when the measurement noise dominates the process noise, $\gamma = 0.5$ and $\beta = 0.1$, and without control weighting.

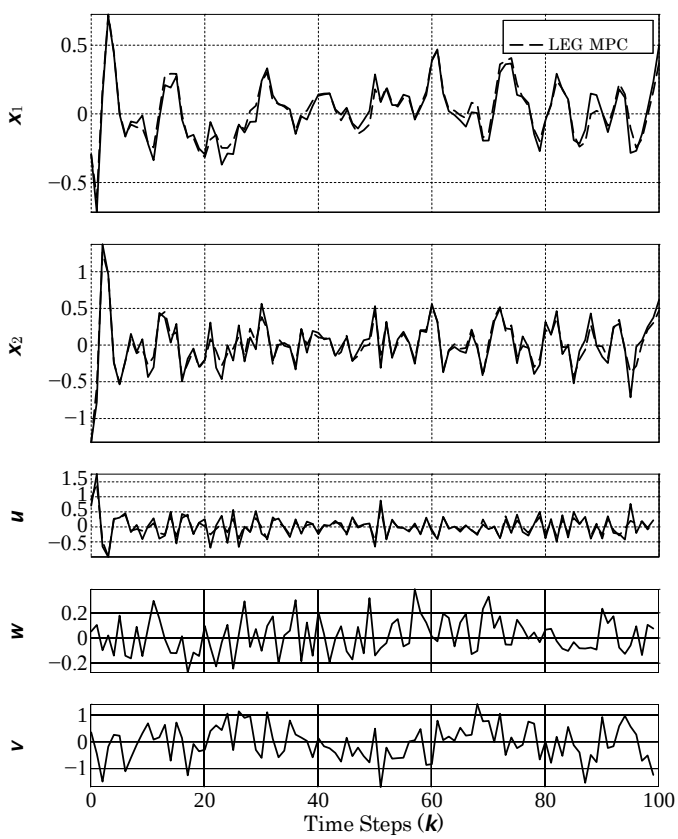


(a) Full view.

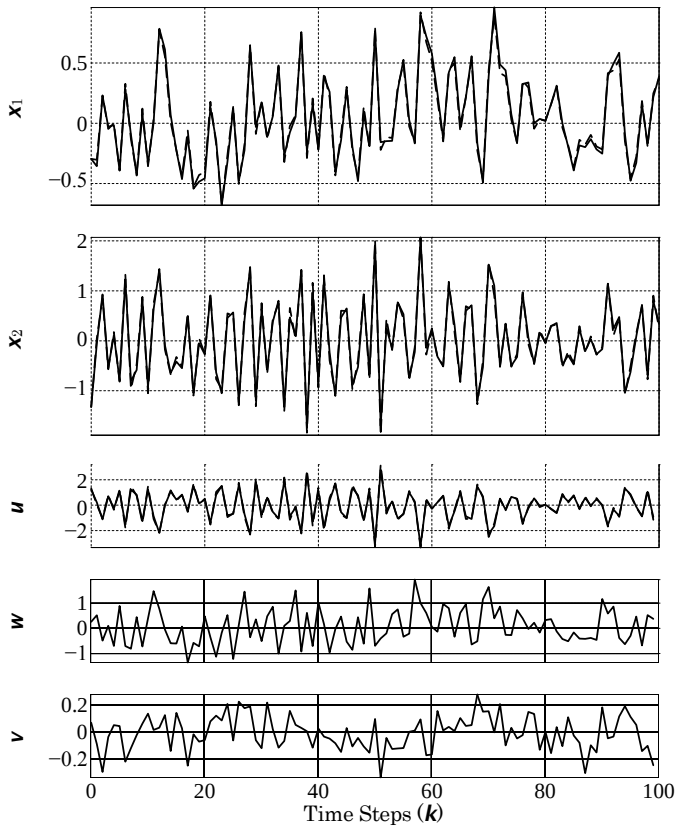


(b) Zoomed view.

Fig. 3. Cauchy and LEG controllers' performance when the measurement noise dominates the process noise, $\gamma = 0.1$ and $\beta = 0.5$, and without control weighting.

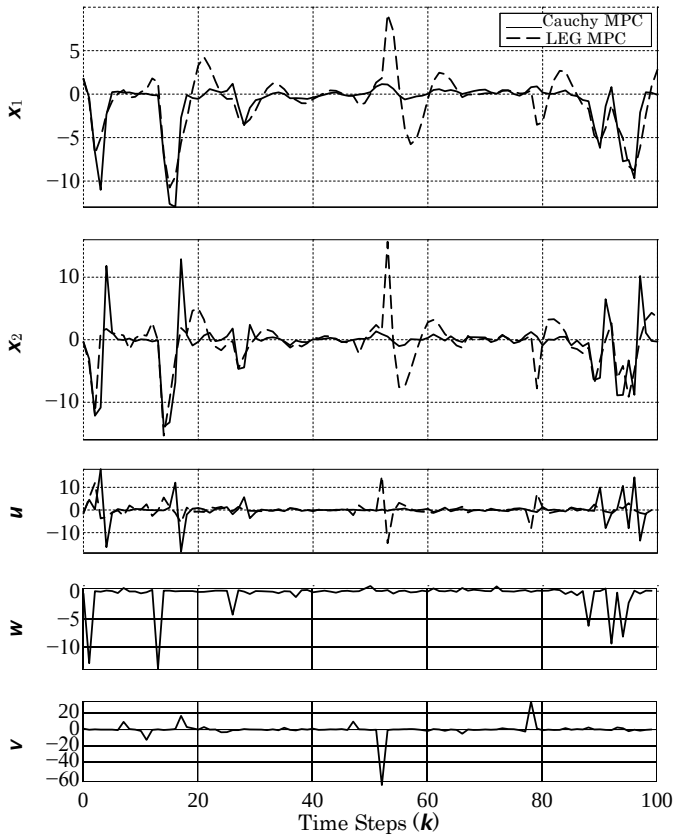


(a) $\gamma = 0.5$, $\beta = 0.1$.

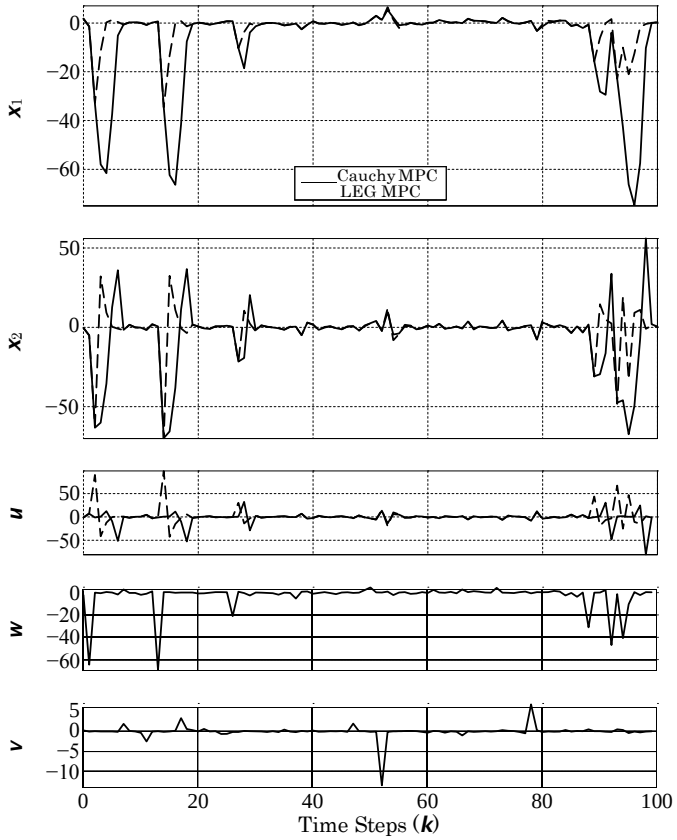


(b) $\gamma = 0.1$, $\beta = 0.5$.

Fig. 4. Cauchy and LEG controllers' performance against Gaussian noises closest in the SaS sense to the given Cauchy parameters.

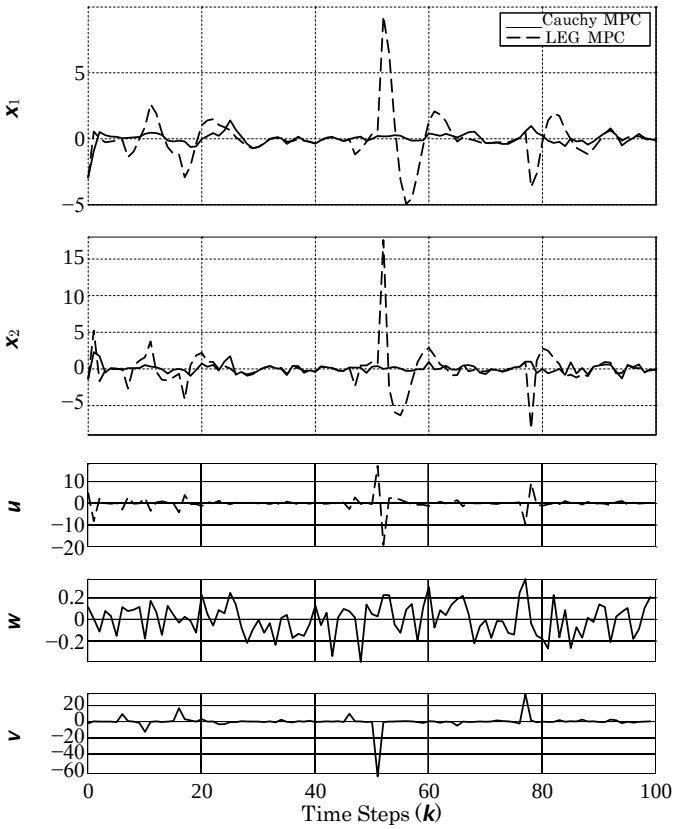


(a) $\gamma = 0.5$, $\beta = 0.1$.

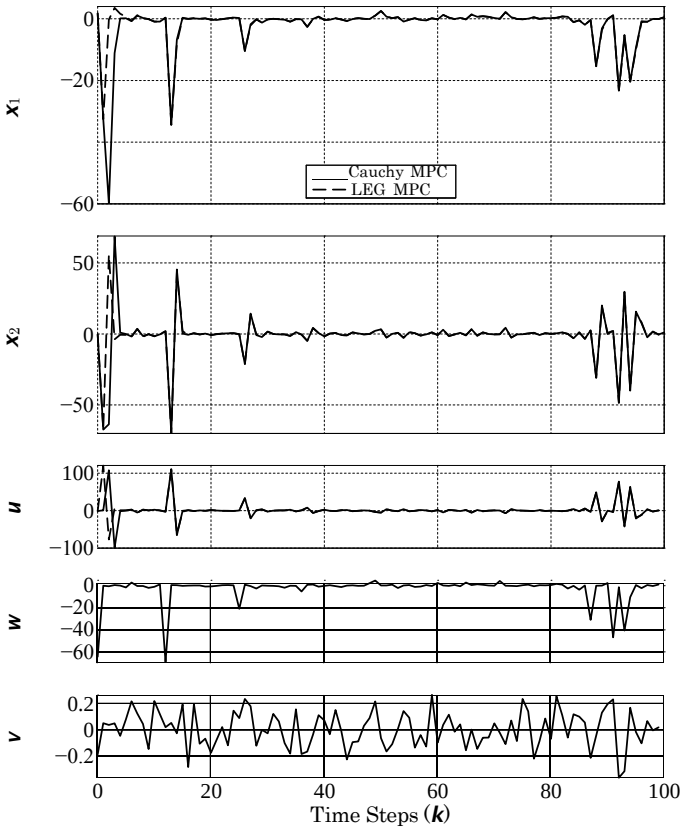


(b) $\gamma = 0.1$, $\beta = 0.5$.

Fig. 5. Cauchy and LEG controllers' performance with control weighting, $\zeta = [10\ 10]$.



(a) Gaussian process noise, Cauchy measurement noise, $\gamma = 0.5$ and $\beta = 0.1$.



(b) Cauchy process noise, Gaussian measurement noise, $\gamma = 0.1$ and $\beta = 0.5$.

Fig. 6. Cauchy and LEG controllers' performance against mixed Cauchy and Gaussian noises.